DOCUMENT RESUME

ED 206 697 TM 810 598

TITLE Testing for College Admissions: Trends and Issues.
INSTITUTION Educational Research Service, Arlington, Va.
PUB DATE 81
NOTE 174p.

EDRS PRICE MF01/PC07 Plus Postage.
DESCRIPTORS Achievement Tests; College Admission; *College Entrance Examinations; Educational Legislation; Higher Education; Scores; Secondary Education; *Standardized Tests; Test Bias; *Test Coaching; Test Construction; *Testing Problems; *Test Reliability; Test Results; *Test Validity
IDENTIFIERS American College Testing Program; College Entrance Examination Board; *Educational Testing Service; Scholastic Aptitude Test; *Truth in Testing

ABSTRACT

Set against the backdrop of a decade in which college admissions test scores have declined, this report reviews issues affecting college admissions testing and their implications, and focuses specifically on the debate between the makers and supporters of standardized tests and test critics. Overviews of the College Entrance Examination Board's (CEEB) Scholastic Aptitude Test (SAT) and the test battery of American College Testing (ACT) are presented and followed by criticisms of: "On Further Examination", a study supported by CEEB to investigate declining test scores. A major portion of the report centers on the debate between consumer advocate Ralph Nader and Educational Testing Service (ETS) concerning the latter organization's status and influence in the field of testing, its accountability, and general issues regarding the reliability and validity of the tests as predictors of students' potential abilities. Reports on the topic of special preparation and coaching for the SAT are reveiwed and criticized. Also documented is the truth-in-testing controversy resulting from the enactment of legislation intended to regulate testing agencies. Two surveys of the use of test scores by colleges and universities are included. In conclusion, a presentation is made of comments from educators on how test scores might be improved. (AEF)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED206697

TM 810 598

Report

Testing for College Admissions: Trends and Issues

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY
C. Randall

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Educational Research Service, Inc.

ERS Report

Testing for College Admissions: Trends and Issues

Copyright © 1981 by Educational Research Service, Inc.
All Rights Reserved

Educational Research Service, Inc.
1800 North Kent Street
Arlington, Virginia 22209
Phone: (703) 243-2100

The Information Source for School Management

Sponsored by

- American Association of School Administrators
- American Association of School Personnel Administrators
- Association of School Business Officials
- Council of Chief State School Officers
- National Association of Elementary School Principals
- National Association of Secondary School Principals
- National School Boards Association
- National School Public Relations Association

Educational Research Service is an independent, nonprofit institution serving the research and information needs of the nation's school systems, their administrative teams, school boards, and the public.

ERS serves as both a national source and clearinghouse for school management research and information. By compiling, analyzing, and sharing information that is essential to effective decision making, ERS assists local school districts in both day-to-day operations and long-range planning. It provides an invaluable approach to data gathering and information reporting through services that are geared to quick yet thorough response to requests from subscribers. An ERS subscription entitles the subscribing school system or agency to: • on-call information service for its administrative staff and governing board • an initial copy of each ERS conducted research study produced during the subscription year • the monthly *ERS Bulletin* summarizing the research of other agencies.

To assure that small as well as large systems can benefit from ERS, a graduated subscription fee has been established. This enables subscribing agencies to share equitably in the cost of compiling and reporting the research needed by all school administrators and governing boards. ERS services are available at reasonable rates to associations of school administrators and school boards, as well as university departments of school administration. Services may be provided to other groups by special action of the ERS Board of Directors. Subscription rates are available upon request.

ERS Executive Staff

Glen E. Robinson
President
Director of Research

John M. Forsyth
Vice President

Prepared by: Joseph E. Garvey

ERS is solely responsible for this publication; no approval or endorsement by specific ERS sponsoring organizations is inferred or implied.

Library of Congress Card Catalog Number: 81-67423

Price of publication: $22.00. To ERS subscribers: single copy mailed routinely as part of the subscription; additional copies, half price.
Stock #219-21652

CONTENTS

Page

LIST OF TABLES iii

LIST OF FIGURES iv

FOREWORD v

ACKNOWLEDGEMENTS vi

INTRODUCTION ix

TRENDS IN COLLEGE ADMISSIONS TEST SCORES 1
  Origins of the College Board's Admissions Examinations 1
  Admissions Test Scores 2
    Scholastic Aptitude Test 2
      Student Descriptive Questionnaire 8
    Test of Standard Written English 9
    Achievement Tests 12
    Preliminary Scholastic Aptitude Test/National Merit Scholarship Qualifying Test 16
    American College Testing Program 19

WHY THE TRENDS IN TEST SCORES 25
  On Further Examination 25
  Criticisms of *On Further Examination* 28
  Recent Theories for Declining Test Scores 29

THE TESTING CONTROVERSY: RALPH NADER VS. ETS 35
  Origins of the Nader-ETS Feud 36
  Major Issues in the Debate 37
    The Nation's Gatekeeper 37
    Monopoly Power 37
    Forced Consumption 38
    Accountability 38
    Nonprofit Status 39
    Center for Occupational and Professional Assessment 40
    ETS Data Banks 40
    Test Development 40
    Standard Error of Measurement 41
    Aptitude 42
    Self-worth 43
    Creativity 44
    Meritocracy 44
    Minorities 45
    Test Scores and Income 47
    Prediction of First Year Grades 48
    Testing the Tests 51

THE COACHING CONTROVERSY: SPECIAL PREPARATION FOR THE SAT VS. NO SPECIAL PREPARATION . . 55
  Commercial Coaching Schools 56
  Boston Regional Office Report on the Effects of Coaching on Standardized Admission Examinations 57
  The Bureau of Consumer Protection's Revised Statistical Analyses of the BRO Report 60
  Criticisms of the BRO Report 63
  ETS's Reanalysis of SAT Data in the 1979 Federal Trade Commission Report 65

Page

NEA's Analysis of SAT Data in the 1979 Federal Trade Commission Report 66
Slack and Porter's Criticisms of ETS 67
Additional Research on Special Preparation Sponsored by the College Board 71
  The Pike and Evans Report 71
  The Pike Report 72
  The Alderman and Powers Report 73
  The Powers and Alderman Report 74
Summary of the Special Preparation Controversy 75

THE TRUTH-IN-TESTING CONTROVERSY: CONSUMER RIGHTS VS. GOVERNMENT INTRUSION 77
Origins of New York's Trust-in-Testing Law 78
Provisions of New York's Truth-in-Testing Law 79
Major Issues Vigorously Debated 80
  Disclosure of Secure Test Questions and Answers 80
  Cost of Test Development 82
  Reduced Services for Handicapped Individuals and Religious Minorities 85
  Errors in Correcting and Reporting Test Scores 86
  Equating the Tests 87
  Disclosure of Background and Statistical Data 87
  Extraterritorial Application 88
Nationwide Movement Predicted 89
Federal Truth-in-Testing 91
Public Interest Principles 93
Summary of the Truth-in-Testing 95

USES OF TEST SCORES BY THE COLLEGES 101
Survey of Undergraduate Admissions Policies, Practices, and Procedures 102
  Importance of Test Scores in Admissions Decisions 102
  Minimum SAT Score Cutoff Points 102
  Freshman Applicants Accepted 102
  SAT Score Expectations 104
  Minimum Grade Point Average Cutoff Point 104
  Credentials Required of Applicants 105
  Exceptions to Formal Admissions Requirements by Selective Colleges 105
  Subjective Judgments of Applicants' Personal Qualities 105
  High School Course Requirements: 1970 vs. 1978 105
Survey of College Admissions Officers 110

HOW TEST SCORES MIGHT BE IMPROVED 115
Perspectives on How Test Scores Might Be Improved 115
A Continuing Controversy: Future Role of Test Scores 125

SUMMARY AND CONCLUSIONS 127

APPENDIXES 139

REFERENCES 147

LIST OF TABLES

Page

1. Percentile Ranks of SAT Scores 3

2. SAT Score Averages for College-Bound Seniors, 1967-1980 4

3. Score Scales, Number of Test Questions and Approximate Incremental Values in Scaled Score for Questions Answered Correctly on the Scholastic Aptitude Test (SAT), Preliminary Scholastic Aptitude Test/National Merit Scholarship Qualifying Test (PSAT/NMSQT), Test of Standard Written English (TSWE), and Achievement Tests, 1980 7

4. Achievement Test Score Averages, 1972-1980 11

5. Reliability Characteristics of the College Board's Achievement Tests 17

6. Verbal Scores for Juniors on the Preliminary Scholastic Aptitude Test/ National Merit Scholarship Qualifying Test 1959-60 to 1979-80 18

7. Mathematical Scores for Juniors on the Preliminary Scholastic Aptitude Test/National Merit Scholarship Qualifying Test 1959-60 to 1979-80 19

8. Mean Scores from Test Section of ACT Assessment Program 22

9. 1973-74 College Bound Seniors Classified by SAT Average and Family Income 49

10. Median Validity Coefficients of Combined-Sex Samples for the SAT-V, SAT-M, High School Record (HSR), and Combined Predictors, 1964-74 50

11. Overall Impact of Coaching: Mean Number of Points (and Confidence Interval) Contributed by Coaching School 62

12. Impact of Coaching Adjusting for Self-Selection: Mean Number of Points (and Confidence Interval) Contributed by Coaching School 63

13. Select Average SAT Point Gain Scores for Coached and Noncoached PSAT and Two-Time SAT Takers for 1975 and 1976 68

14. Effects of Coaching on SAT Scores in Studies Cited and Not Cited by Educational Testing Service 69

15. Comparative Analysis of Program Costs by Activity, 1970-71 83

16. Admissions Testing Program Revenues and Expenses 83

17. ETS Test Development Spending Breakdown and Budget 84

18. The Debate over Testing Legislation: Pros and Cons 96

19. Importance of Admissions Test Scores in Decision Making, by Type of Institution . . 103

20. Minimum SAT Scores Below Which Applicants Generally Are Not Considered Eligible for Admission, by Type of Institution 103

21. Percent of Freshmen Applicants Accepted and Enrolled, by Type and Control of Institution 104

Page

22. Credentials Required of All Applicants, by Type and Control of Institution 106

23. Exceptions to the Formal Requirements for Admission, Selective Institutions, by Type of Institution 107

24. Percent of Respondents Who Would Consider Subjective Judgments of Personal Qualities, by Type of Institution 108

25. High School Course Requirements in 1970 and 1978, by Type of Institution 109

LIST OF FIGURES

1. SAT Verbal Score Averages for College-Bound Seniors, 1967-1980 5

2. SAT Mathematical Score Averages for College-Bound Seniors, 1967-1980 6

3. Test of Standard Written English Averages by Sex, 1975-1980 10

4. Achievement Test Score Averages, 1972-1980: English Composition, Literature, American History/Social Studies, and European History and World Cultures 12

5. Achievement Test Score Averages, 1972-1980: Math Level I, Math Level II, Biology, Chemistry, and Physics 13

6. Achievement Test Score Averages, 1972-1980: French, German, Hebrew, Latin, Russian, and Spanish 15

7. Preliminary Scholastic Aptitude Test/National Merit Scholarship Qualifying Test Means of Mathematical and Verbal Scores for Male and Female Juniors Tested from 1959-1960 to 1979-1980 16

8. Mean Scores from Test Section of ACT Assessment Program: Males, Females, and Males/Females Combined, 1969-1970 to 1979-1980 20

8

FOREWORD

For more than a decade declining college admissions test scores have posted serious questions for school administrators, board members, teachers, and the public. Standardized college admissions tests such as the College Entrance Examination Board's Scholastic Aptitude Test and the American College Testing Program's battery of tests are often viewed as barometers of school and program effectiveness and as indicators of a person's native intelligence, although these tests were [illegible]t designed to fulfill these functions.

How serious is the decline in college admissions test scores? How can the downward trend in SAT scores be reversed? What are the implications for students, schools, and communities? These are some of the hard questions that school boards and school policy makers have attempted to answer since the mid-1970s. In some instances, intensified inquiry into the transition from high school to college has resulted in a reorganization of the high school curriculum, a mandate for minimum competency testing, and an expanding debate over the role of standardized testing in American society. Public attention is now focused on the debate between the makers of standardized tests and numerous test critics including consumer advocates, public interest groups, and some educator organizations. Standardized tests such as the SAT and ACT are the object of intense scrutiny to determine whether they are, in fact, valid, reliable, and able to predict a person's potential to perform college-level work.

This ERS Report provides a comprehensive review of recent trends in college admissions test scores, as well as an objective analysis of major issues affecting college testing. The Report summarizes the debate between consumer advocate Ralph Nader and the Educational Testing Service, the issue of special preparation for the SAT, and recently enacted and proposed truth-in-testing legislation that would regulate the testing agencies. Also included are two surveys that provide valuable information concerning the use of test scores by colleges and universities. Statements from well-known educators present a variety of perspectives on how test scores might be improved. Some *tentative* conclusions based on the data and information currently available are presented.

We hope that this Report will be useful to school administrators, board members, teachers and others who must deal with the numerous and highly complex issues affecting testing for college admissions.

Glen Robinson
Director of Research
Educational Research Service

ACKNOWLEDGEMENTS

ERS appreciates the many contributions made by educational organizations, professional educators, and concerned individuals in the preparation of this Report. ERS gratefully acknowledges permission to use excerpts from the following publications:

"ACT Assessment Test Scores--3rd Year in Row for Increase," *Activity*, May 1980, p. 4. Copyright 1980 by The American College Testing Program, Iowa City, IA.

An Agenda for Action: Recommendations for School Mathematics of the 1980s. Copyright 1980 by the National Council of Teachers of Mathematics, Reston, VA.

ATP Guide for High Schools and Colleges 1979-81. Copyright 1979, 1980 by the College Entrance Examination Board, New York, NY.

Guide to the Admissions Testing Program 1978-79. Copyright 1978 by the College Entrance Examination Board, New York, NY.

"Joint Statement on Principles of Good Practice in College Admissions and Recruitment." Published in "Self-Regulation Initiatives: Guidelines for Colleges and Universities," October 1979, p. 2, by the American Council on Education, Washington, DC.

Measurement and Testing: An NEA Perspective. Copyright 1980 by the National Education Association, Washington, DC.

National College-Bound Seniors, 1980. Copyright 1980 by the College Entrance Examination Board, New York, NY.

National College-Bound Seniors, 1979. Copyright 1979 by the College Entrance Examination Board, New York, NY.

National Report College-Bound Seniors, 1978, 1977, 1975-76. Copyright 1978, 1977, 1976, 1975 by the College Entrance Examination Board, New York, NY.

"National Testing Movement." Basic Policy Statement, Resolutions on Education. Adopted in 1979 by the National Association for the Advancement of Colored People, Washington, DC.

"Recommendations for the Preparation of High School Students for College Mathematics Courses." Published by the Mathematical Association of America, Washington, DC.

"The Scholastic Aptitude Test: A Critical Appraisal," by Warner V. Slack and Douglas Porter. Published in the May 1980 issue of the *Harvard Educational Review*, pp. 154-175. Copyright 1980 by the President and Fellows of Harvard College.

Searching for the Truth about "Truth-in-Testing" Legislation: A Background Report, Report No. 132, by Rexford Brown and Merle Steven McClung. Published in January 1980 by the Education Commission of the States, Denver, CO.

"Statement of Factual Correction and Rebuttal to Educational Testing Service Memorandum in Opposition," May 18, 1979, by New York Public Interest Research Group, Albany, NY.

Test Scores and Family Income. Copyright 1980 by the Educational Testing Service, Princeton, NJ.

Test Use and Validity. Copyright 1980 by the Educational Testing Service, Princeton, NJ.

ERS would also like to extend special appreciation to the following individuals for their contributions to this publication: Rexford Brown, Director of Publications and User Products, National Assessment of Educational Progress; Hugh L. Carey, Governor of the State of New York; Roger Farr, Professor of Education at Indiana University and President of the International Reading Association; George H. Hanford, President of the College Entrance Examination Board; Shirley A. Hill, President of the National Council of Teachers of Mathematics; Stanley H. Kaplan, Founder and President of Stanley H. Kaplan Educational Center, Ltd.; Nancy Kokus, Guidance Project Specialist, Fairfax County Public Schools, Fairfax, VA; Kenneth P. LaValle, New York State Senator, (R); Merle Steven McClung, Director of the Education Commission of the States Law and Education Center; Willard H. McGuire, President of the National Education Association; Ralph Nader, Consumer Advocate; Allan Nairn, Consumer Advocate; Vito Perrone, Dean of the Center for Teaching and Learning, the University of North Dakota; Lewis W. Pike, National Institute of Education; Alan C. Purves, Principal of University High School at the University of Illinois and President of the National Council of Teachers of English; Diane Ravitch, Associate Professor of History and Education, Teachers College, Columbia University; Philip R. Rever, Director of the Washington, DC, Office of the American College Testing Program; Scott D. Thomson, Executive Director of the National Association of Secondary School Principals; Shelia Tobias, Member of the Board of Curators at Stevens College and author of *Overcoming Math Anxiety*; William W. Turnbull, former President of the Educational Testing Service; James G. Ward, Director of Research, the American Federation of Teachers; and Ted Weiss, U.S. Representative (D-N.Y.).

INTRODUCTION

The continuing national controversy over standardized testing affects American education from elementary schools to institutions of higher education and confronts educators, school officials and others with difficult questions. For nearly two decades many educators and much of the public have been concerned and alarmed by reports of declining college admissions test scores. Broad media attention to declining college admissions test scores has resulted in public concern and demand for immediate remedies for what many perceive to be a decline in educational quality throughout the nation.

With the recent emphasis on basic educational skills and the tightening of academic standards, many persons have been hoping for a reversal in declining standardized college admissions examination scores. However, scores on the Scholastic Aptitude Test (SAT), taken by one-third of the nation's three million high school seniors, dropped to record low levels in 1980, continuing a trend that began more than 15 years ago. [69:3; 6.2:4; 289.1]* This decline has been viewed by some as evidence of deterioration of the nation's secondary schools and has helped spark nationwide campaigns to regulate standardized testing. [31:A3; 123:3]

The continued test score decline has stimulated serious debate over whether the SAT is an appropriate measure of educational quality, whether other measures can better indicate intellectual readiness for academic work and predict college potential, whether the SAT discriminates against minority students and perpetuates a class system "in the guise of merit," and whether the test makers are accountable to the public. [187:58G-64GE; 188:219]

The purposes of this ERS Report are to assemble facts about recent trends in college admission testing among secondary school students, collect information and opinions regarding the reasons for declining college admissions test scores, and summarize suggestions for improving test scores. In addition, important issues affecting college admissions are examined. Included are summaries of the debate between Ralph Nader and the Educational Testing Service (ETS), the effects of coaching in preparation for the SAT, and the background of truth-in-testing legislation. Statements from college admissions personnel on the use of test scores are also provided to give an overview of the college admissions process.

**Bibliographic references cited in this text are noted by numbers within brackets. The number before the colon indicates the entry number within the reference section beginning on page 147; the number following the colon indicates the page(s) within the entry. Where no colon appears the citation refers to the entire entry. Multiple citations are separated by semicolons.*

TRENDS IN COLLEGE ADMISSIONS TEST SCORES

Secondary schools across the nation are under fire to reverse the test score decline that is well into its second decade. [168:E11] In particular, attention has centered on the Scholastic Aptitude Test and the American College Testing Program (ACT). The SAT and ACT are standardized college admissions tests designed to measure a candidate's developed mental abilities and potential to perform first-year college work. These tests are intended to supplement the candidate's academic record and accomplishments. [19:15; 17:1, 3] Since the quality of the high school curricula and standards in grading practices vary considerably among the nation's high schools, standardized admissions tests such as the SAT and ACT are used to provide a "common currency" that allows admissions officers and counselors to place students on the same footing regardless of social strata and geographic location. [69:2-3] Colleges differ, however, in the weight given to admissions test scores in the selection of freshmen. Many private four-year colleges consider test scores a very important factor in the admissions process, while many two-year colleges do not require them. [44:4]

Origins of the College Board's Admissions Examinations

In an attempt to introduce order into the transition from high school to college, the College Entrance Examination Board (CEEB) was organized at Columbia University in November 1900. The first CEEB admissions tests were essay examinations designed by well-known scholars. In 1901, 973 candidates wrote essays in history, Greek, Latin, German, French, English, mathematics, chemistry, and physics. In 1926, Carl Campbell Brigham, a Princeton University psychologist, introduced the more familiar multiple-choice SAT. The original multiple-choice SAT, administered to 8,040 candidates, consisted of nine subtests: paragraph reading, logical inference, analogies, definitions, artificial language, antonyms, arithmetical problems, classification, and number series. Three years later Brigham divided the SAT into two separate sections measuring verbal and mathematical aptitude. During the 1930s CEEB concentrated on improving the consistency of its operations and strengthening the technical aspects of test construction. The increasing numbers of candidates taking the SAT prompted CEEB to provide a means of comparing SAT scores among the different standardized test forms. In April 1941, the group of 10,654 high school seniors tested became the standardization group for all subsequent forms of the SAT. Since then, SAT scores have been equated directly to preceding test forms and indirectly to the April 1941 standardized form. This procedure, according to CEEB, insures that test scores have the same meaning from year to year and that the scoring represents the same level of ability regardless of the group tested, the difficulty of the test, or the time of year tested. [19:1-3, 32]

The SAT is designed to measure verbal and mathematical reasoning abilities on a standardized score scale ranging from 200 to 800. A score of 200 represents the lowest score reported, while a score of 800 represents the highest score reported. Thus, the 500-point merely marks the midway between the two end points of the scale. In April 1941 the mean score and standard deviation of the initial standardization group were 500 and 100, respectively. These scores, however, do not represent the mean score and standard deviation of all high school seniors or of all college freshmen who have since taken the SAT. [19:11] Since the SAT candidate group is self-selected and does not represent all high school seniors who plan to attend college, special norm studies were conducted in 1960, 1966, and 1974 to obtain representative norm data on all high school seniors. [19:81-82; 6:15-16]

As shown in Table 1, the mean norm SAT score for the national high school sample is 368 for the verbal section and 402 for the mathematical section. Norms for college-bound seniors are 61 points higher on the verbal section and 66 points higher on the mathematical section. For the math section alone, men in the national high school sample averaged 26 points higher than women in the national sample; of college-bound seniors, men averaged 50 points higher than women.

Admissions Test Scores

Scholastic Aptitude Test.— Each year approximately one million high school seniors seeking admission to college take the SAT. While the College Entrance Examination Board located in New York City determines the policies and is responsible for the SAT, the Educational Testing Service of Princeton, New Jersey, develops and administers the tests, reports and analyzes the test scores, conducts educational research, and gives instructional and advisory services in measurement and testing, both nationwide and abroad. [6.2:4; 69:5]

Today the SAT is a 2½ hour multiple-choice test consisting of two 30-minute verbal sections with a total of 85 questions; two 30-minute mathematical sections with a total of 60 questions; and one 30-minute experimental section consisting of questions used to equate new editions of the test to old versions and to improve test quality. The two verbal sections (SAT-V) contain antonym questions to test the extent of a candidate's vocabulary, analogy questions to test the ability to recognize a relationship in a pair of words, sentence completion questions to test the ability to recognize relationships among the parts of a sentence, and reading passages to test interpretive and analytical skills. The antonym and analogy questions comprise the vocabulary subscore and the sentence completion and reading passage questions comprise the reading subscore. The mathematical section (SAT-M) includes arithmetic, algebra, and geometry in two formats: 40 standard multiple-choice questions and 20 quantitative comparisons. [6.1:4, 8-9]

In 1962-63, test scores climbed to a peak SAT-V average of 478 and a peak SAT-M average of 502. Since then, both score averages have dropped substantially, especially in the verbal section. [297:6] In response to the unprecedented decline in test scores, S. P. Marland, Jr., then president of the College Board, appointed a "blue-ribbon" Advisory Panel in 1975 to investigate the test score decline and interrelated issues. Willard Wirtz, former Secretary of Labor, chaired the twenty-one member Advisory Panel on the Scholastic Aptitude Test Score Decline. The Panel's summary report, *On Further Examination*, was released in 1977. As the Advisory Panel investigated possible reasons

TABLE 1.--Percentile Ranks of SAT Scores

	For national high school sample and for college-bound seniors*					For men and women separately on the mathematical sections			
	Verbal		Mathematical			Men		Women	
Score	National high school sample	College-bound seniors	National high school sample	College-bound seniors	Score	National high school sample	College-bound seniors	National high school sample	College-bound seniors
800	99+	99+	99+	99+	800	99+	99+	99+	99+
750	99+	99+	99+	99	750	99+	98	99+	99+
700	99+	99	99	97	700	99	95	99	99
650	99	97	98	92	650	96	88	98	96
600	97	92	94	84	600	92	78	96	90
550	93	85	88	74	550	85	66	91	81
500	86	73	80	59	500	75	50	84	68
450	76	57	69	44	450	64	36	73	52
400	64	40	56	30	400	52	23	61	36
350	49	24	41	17	350	38	13	46	21
300	33	11	23	6	300	20	5	25	8
250	17	4	5	1	250	4	1	6	1
200	--	--	--	--	200	--	--	--	--
Mean	368	429	402	468	Mean	416	494	390	444
Standard deviation	111	110	112	118	Standard deviation	117	121	104	110

*Percentile ranks for the national high school sample are based on the scores earned by a representative sample of high school students who took the test at a special administration in October 1974. Students in this sample were selected regardless of course of study or intention to enroll in college. Percentile ranks for college-bound seniors are based on the most recent scores earned by students in the 1978 graduating class who took the SAT at any time while in high school. The percentile ranks for the SAT-Verbal and SAT-Mathematical scores are found on ATP reports, while the percentile ranks of SAT-Mathematical scores for men and women separately are *not* used on ATP reports. (Emphasis in the original)

SOURCE: Admissions Testing Program of the College Board. *ATP Guide for High Schools and Colleges 1979-81*, Princeton, N.J.: College Entrance Examination Board, 1979, 1980, pp. 15-16. Copyright 1979, 1980 by College Entrance Examination Board. Used with permission.

TABLE 2.--SAT Score Averages for College-Bound Seniors, 1967-1980*

Verbal Section					
Testing Year	Male	Female	Total	Annual Raw Score Change	Annual Percent Change
1967	463	468	466	---	---
1968	464	466	466	0	---
1969	459	466	463	-3	-0.6
1970	459	461	460	-3	-0.6
1971	454	457	455	-5	-1.1
1972	454	452	453	-2	-0.4
1973	446	443	445	-8	-1.8
1974	447	442	444	-1	-0.2
1975	437	431	434	-10	-2.2
1976	433	430	431	-3	-0.7
1977	431	427	429	-2	-0.5
1978	433	425	429	0	---
1979	431	423	427	-2	-0.5
1980	428	420	424	-3	-0.7

Mathematical Section					
Testing Year	Male	Female	Total	Annual Raw Score Change	Annual Percent Change
1967	514	467	492	---	---
1968	512	470	492	0	---
1969	513	470	493	+1	+0.2
1970	509	465	488	-5	-1.0
1971	507	466	488	0	---
1972	505	461	484	-4	-0.8
1973	502	460	481	-3	-0.6
1974	501	459	480	-1	-0.2
1975	495	449	472	-8	-1.7
1976	497	447	472	0	---
1977	497	445	470	-2	-0.4
1978	494	444	468	-2	-0.4
1979	493	443	467	-1	-0.2
1980	491	443	466	-1	-0.2

*The averages for 1967 through 1971 are estimates of the averages that would have been reported for college-bound seniors of those years if such reports had been produced.

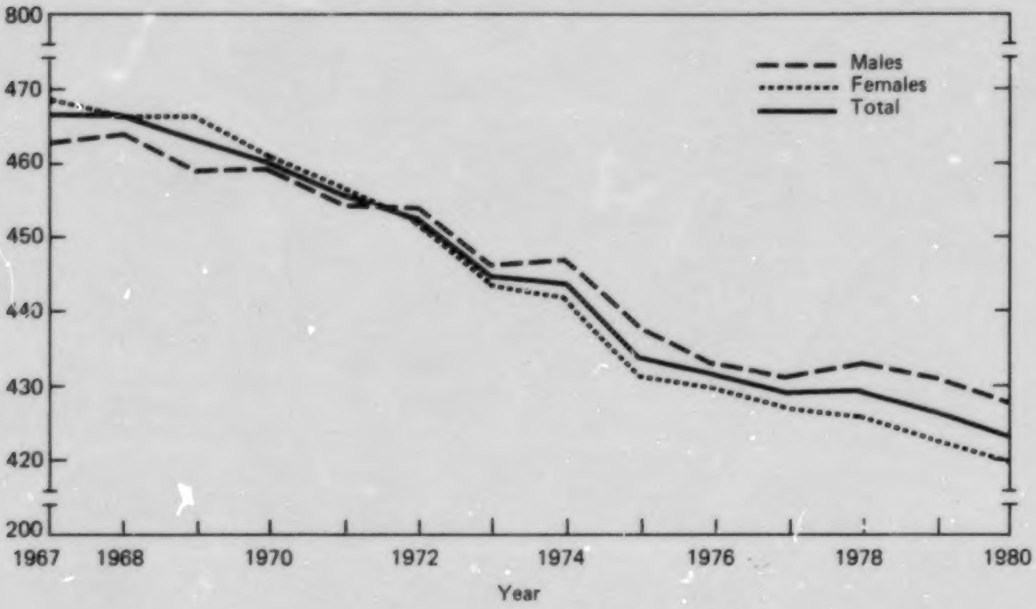
SOURCE: Admissions Testing Program of the College Board. *National College-Bound Seniors, 1980.* New York, NY: CEEB, 1980, p. 5. Copyright 1980 by College Entrance Examination Board. Used with permission.

for the decline in SAT scores, the SAT verbal and mathematical test scores continued their downward spiral.

Table 2 lists the national SAT-V and SAT-M score averages for male and female college-bound seniors from 1967 to 1980. During this period, average SAT-V scores have declined 42 points for all college-bound seniors, 35 points for males, and 48 points for females. Average SAT-M scores have fallen 26 points for all college-bound seniors, 23 points for males, and 24 points for females.

Since 1976 the mean SAT-V score average for college-bound seniors had declined seven points,

FIGURE 1.--SAT Verbal Score Averages for College-Bound Seniors, 1967-1980*



*The averages for 1967 through 1971 are estimates of the averages that would have been reported for college-bound seniors of those years if such reports had been produced. Each of the 85 SAT-V questions is worth approximately seven test points on the College Board's 200-to-800 point scale.

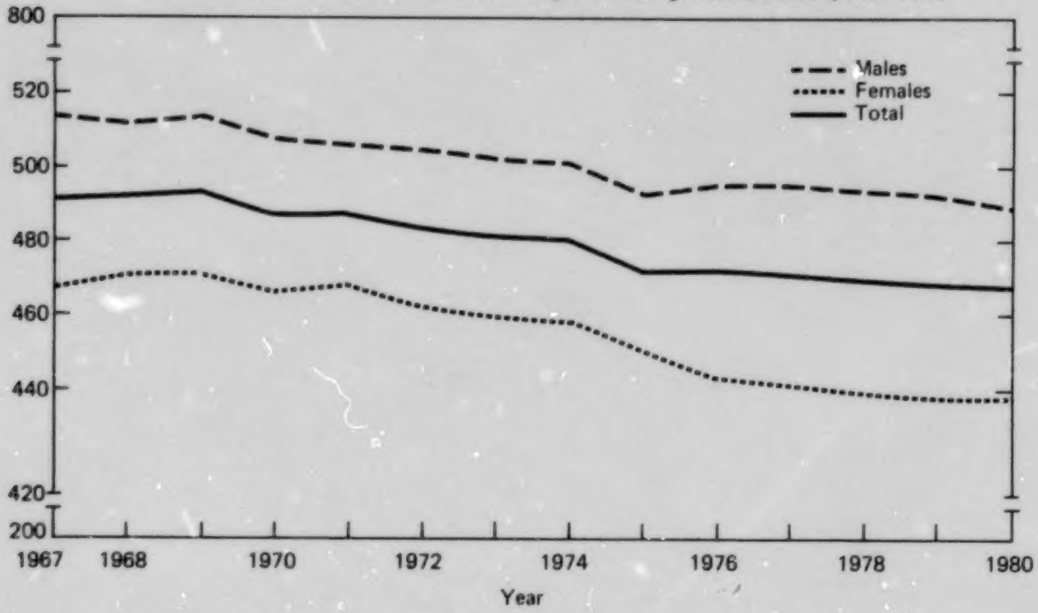
SOURCE: Admissions Testing Program of the College Board. *National College-Bound Seniors, 1980.* New York, NY: CEEB, 1980, p. 5. Copyright 1980 by College Entrance Examination Board. Used with permission.

---

from 431 to 424. The mean SAT-M score average has declined six points, from 472 to 466. Between 1976 and 1977 both SAT-V and SAT-M score averages declined two points. From 1977 to 1978 SAT-V score averages stabilized at 429. This temporary leveling off was the first positive sign in SAT-V scores since 1968. During the 1977-78 period, however, the SAT-M score averages declined two points, from 470 to 468. Between 1978 and 1979 the verbal score average declined two points, and the mathematical score average, one point. In 1980 the verbal score average declined three points, to 424, and the mathematical score average declined one point to 466. Thus, the 1980 SAT score averages represent the lowest recorded levels on the SAT. [6.2:5] CEEB characterized the most recent declines as disappointing in light of "efforts by many schools to improve education." [289.1:4]

The SAT-V score average of males fluctuated between 433 and 431 between 1976 and 1979. In 1980 the SAT-V score average of males declined three points from 431 to 428. Since 1976, females' verbal score average has declined consistently from 430 to 420. Between 1976 and 1977 the SAT-M score average of males stablized at 497, and then declined three points in 1978. The following year the SAT-M score average of males declined another point. In 1980 the score average of males declined two additional points to 491.

FIGURE 2.--SAT Mathematical Score Averages for College-Bound Seniors, 1967-1980*



*The averages for 1967 through 1971 are estimates of the averages that would have been reported for college-bound seniors of those years if such reports had been produced. Each of the 60 SAT-M questions is worth approximately nine test points on the College Board's 200-to-800 point scale.

SOURCE: Admissions Testing Program of the College Board. *National College-Bound Seniors, 1980.* New York, NY; CEEB, 1980, p. 5. Copyright 1980 by College Entrance Examination Board. Used with permission.

From 1976 to 1979 the SAT-M score average of females showed an annual decline of one point. In 1980 the score average of females stabilized at 443. [6.2:5]

Table 2 also shows that between 1968 and 1972 score averages for college-bound seniors declined continually on the SAT-V and fluctuated on the SAT-M. Between 1973 and 1975 the decreases in both SAT tests became more acute. Since 1976 SAT score averages have continued to decline, but at a lesser rate than in the early 1970s.

The downward trend exhibited by the SAT-V score averages for college-bound seniors from 1967 to 1980 is indicated in Figure 1. The trend on the SAT-M section is shown in Figure 2.

Table 3 shows the approximate values of the individual ATP tests, the number of test questions on each test, and the range of their score scales. According to ETS, the 85 SAT-V test items are worth approximately seven points each and the 60 SAT-M test items are worth approximately nine points each on the College Board's 200-to-800 point scale. These point values vary slightly depending on the difficulty and length of the test items. Between 1967 and 1980 the combined score averages for males and females on the SAT-V declined 42 points (an average annual decrease of 3.23 test points). Applying ETS's estimated value of seven points for

TABLE 3.--Score Scales, Number of Test Questions and Approximate Incremental Values in Scaled Score for Questions Answered Correctly on the Scholastic Aptitude Test (SAT), Preliminary Scholastic Aptitude Test/National Merit Scholarship Qualifying Test (PSAT/NMSQT), Test of Standard Written English (TSWE), and Achievement Tests, 1980

Individual ATP Test	Score Scale	Number of Test Questions	Approximate Values
1. SAT-Verbal	200-800	85	7
2. SAT-Mathematical	200-800	60	9
3. PSAT-Verbal	20-80	65	0.9
4. PSAT-Mathematical	20-80	50	1.0
5. TSWE	20-80 Maximum score reported is 60+	50	1.0
6. English Composition	200-800	95 on old forms; 90 on new forms; 70 when the ECT has an essay component	6
7. Literature	200-800	60	9
8. American History and Social Studies	200-800	100	6
9. European History and World Cultures	200-800	100	6
10. Mathematics Level I	200-800	50	10
11. Mathematics Level II	200-800	50	9
12. Biology	200-800	100	6
13. Chemistry	200-800	90	6
14. Physics	200-800	75	6
15. French	200-800	85	6
16. Spanish	200-800	85	5
17. German	200-800	80	6
18. Latin	200-800	75	6
19. Hebrew	200-800	90	4
20. Russian	200-800	80	6

*Approximate incremental value in scaled score contributed by each question answered correctly. Four decimal values are used in the actual conversion of raw scores to the College Board Scale. Estimated incremental values may vary slightly from one form of the test to another because of the differences in test difficulty, length of reading passages, or occasional variations in the number of test questions.

SOURCE: Information used by permission of Educational Testing Service, Princeton, NJ, December 1980.

each SAT-V test item indicates that 1980 test takers answered approximately six fewer verbal test items correctly compared to their counterparts in 1967. This represents an average annual decrease of 0.46 SAT-V test items; that is, every two years the average SAT taker missed one additional verbal test question.

Between 1967 and 1980 combined score averages for males and females on the SAT-M declined 26 points (an average annual decrease of 2.00 test points). Likewise, applying ETS's estimated value of nine points for each SAT-M test item shows that, in 1980, SAT test takers answered 2.88 fewer mathematical test items correctly compared to the average 1967 test taker. This represents an average annual decrease of 0.22 SAT-M test items. Thus, every four years the average SAT taker missed one extra mathematical test question.

The combined total SAT-V and SAT-M score averages for males and females declined 68 points between 1967 and 1980. This is equal to an average annual decline of about 5.23 test points. At approximately eight points per question, this means that 1980 SAT test takers answered approximately 8.50 fewer test items correctly on a 145-question exam compared to 1967 SAT test takers. Over the 13-year period, this corresponds to a total decrease of two-thirds of a test item each year for the average SAT test taker.

Since 1976 the number of seniors with SAT scores above 600 has decreased from eight percent to seven percent on the verbal section and from 17 percent to 15 percent on the mathematical section. The number of candidates scoring in the middle range, between 400 and 600, has declined two percent on the verbal section to 50 percent and declined one percent to 54 percent on the mathematical section. By 1980 the number of students scoring below 400 increased two percent in the verbal section, to 42 percent, and increased one percent in the mathematical section, to 30 percent. [10:14; 6.2:12]

Traditionally, more male candidates than female candidates have taken the SAT. The 1974-75 academic year marked the first time that the number of females exceeded males. In 1980, for the sixth consecutive year, the percentage of female test takers exceeded males, 51.8 percent to 48.2 percent. [10:5; 6.2:4]

According to the College Board, the scores of two individuals on the same SAT must differ by more than 1.5 times the standard error of the difference to ensure that the higher score indicates greater ability and is not the result of measurement error alone. Latest data for the SAT (Form 3, administered in May 1978) shows that the standard error of the difference for the SAT-V is 43, and for the SAT-M, 46. [6:10] This means that two SAT-V scores would have to differ by more than 65 points (43 x 1.5) and two SAT-M scores would have to differ by more than 69 points (46 x 1.5) for one to be sure that "there is a genuine difference in the abilities being measured." [6:11] Differences in test scores of two individuals of *less* than 1.5 times the standard error of the difference should *not* be considered significant, the College Board states.

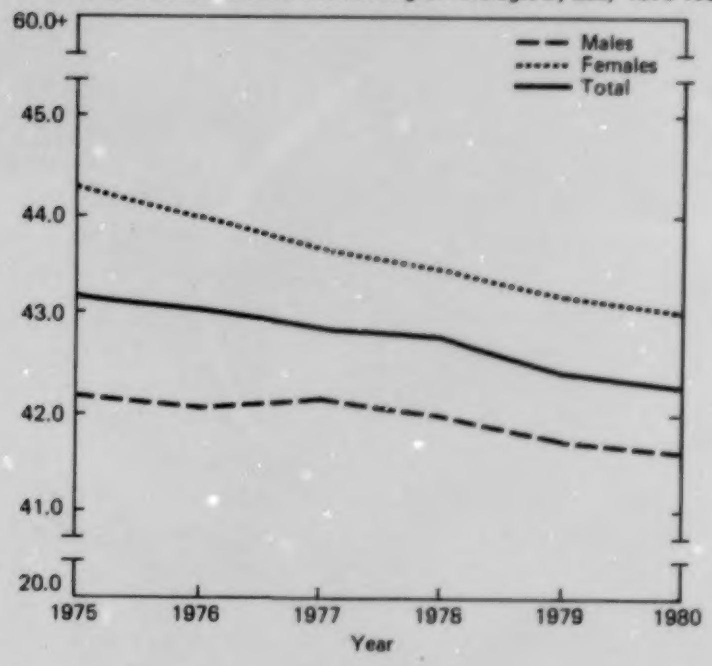
Student Descriptive Questionnaire.-- Each year students who take the SAT are asked to complete the Student Descriptive Questionnaire (SDQ) which includes questions about the candidate's background, academic accomplishments, extracurricula activities, ethnic group, family income, and intended area of college study. The voluntary response rate of the SDQ has risen from 77 percent in 1976 to 91 percent in 1980, thereby providing a nearly complete profile of about two-thirds of all high school graduates who go immediately to college. [6.2:4] A comparison of the responses of the high school seniors on the 1976 and 1980 SDQs indicates the following trends:

- The percentage of students who identified themselves as belonging to an *ethnic minority* rose from 15 percent in 1976 to an all-time high of 18 percent in 1980. Women comprised 55 percent of all minority students. [10:5; 6.2:7] Recently released data (1980) comparing the test score averages of 350,000 whites, 31,000 blacks, and 5,500 Chicanos who took the SAT between 1972-73 and 1976-77 indicated that whites averaged 455 on the SAT-V and 493 on the SAT-M. In contrast, Chicanos' averages were 80 SAT-V points and 81 SAT-M points lower than whites, while blacks' averages were 119 SAT-V points and 134 SAT-M points lower than whites. In testimony before the House of Representatives Subcommittee on the Civil Service, Winton H. Manning, a senior vice-president of ETS, stated that differences in test scores among ethnic groups should "come as no surprise to anyone familiar with the historically unequal education available to blacks as compared with whites or with corresponding differences in social, economic, and occupational spheres of American life." [135:5]

- Between 1976 and 1980 the median *parental contribution to the costs of education* declined, from $1,170 to $920. The ability of families to contribute financially varied considerably among ethnic groups. Since 1976 the median contribution of whites declined from $1,380 to $1,160; for Orientals, from $860 to $520; for American Indians, from $570 to $440; for Mexican Americans, from $210 to $50; for Puerto Ricans, from $200 to $0; and for blacks, from $100 to $0. About half of Puerto Rican and black families could not make financial contributions, while the other half could make some contribution. The College Board also reported in 1980 that 21 percent of families could pay for the entire annual costs of education at a public four-year college (averaging $3,409), while 11 percent of families could afford the average annual cost of a private four-year college ($6,082). [10:6; 6.2:8]
- The percentage of seniors indicating that they planned to seek *part-time work in college* to help finance their education declined from 43 percent in 1976 to 39 percent in 1980. [10:6; 6.2:9]
- For the fourth consecutive year, the *estimated grade point average* (GPA) for males and females combined declined from 3.12 to 3.06 on a four-point scale. In 1980 males reported GPAs of 3.00 compared to 3.05 in 1976; females reported GPAs of 3.12 compared to 3.18. During this period, the percentage of students who indicated that they had taken advanced or honors courses in English declined by one percent (to 14 percent), and in mathematics by two percent (to 14 percent). [10:6; 6.2:6-7]
- In 1980, 71 percent of the seniors reported themselves to be in the top two-fifths of their *high school class* compared to 75 percent in 1976. Almost all the other seniors reported themselves to be in the middle fifth in both years. [10:6; 6.2:7]
- Since 1976 the gap between the percentage of men and women *aspiring toward a professional degree* has narrowed. By 1980 the percentage of women reporting plans to study for a professional degree rose from 15 percent to 16 percent, while the percentage of men declined from 24 percent to 20 percent. [10:9; 6.2:8-9]
- The 1980 seniors indicated that their most popular *area of intended study* was business and commerce, up from 12.6 percent in 1976 to 18.6 percent in 1980. Interest in health and medicine declined slightly to second place, from 17.9 percent to 14.7 percent. Engineering registered gains from 8.4 percent to 11.1 percent, and the social sciences, from 6.8 percent to 7.8 percent. Continuing a trend that began in 1973, interest in education declined from 8.7 percent to 6.1 percent. [10:10; 6.2:9]

Test of Standard Written English.-- The Test of Standard Written English (TSWE) is administered with the SAT. The 30-minute TSWE contains 50 multiple-choice questions designed to evaluate a candidate's ability to recognize the basic principles of grammar, usage, diction, punctuation, and effectiveness of expression. [6.1:9] The TSWE has 35 usage questions and 15 sentence correction questions. [11:16] Scores are reported on a 20-to-80 scale, which corresponds to the SAT scale of 200 to 800. Since the TSWE is intended only to assist college admissions personnel in placing students in appropriate

FIGURE 3.--Test of Standard Written English Averages by Sex, 1975-1980*



*Each of the 50 TSWE questions is worth approximately 1.0 on the College Board's scale of 20 to 80.

SOURCE: Admissions Testing Program of the College Board. *National College-Bound Seniors, 1980.* New York, NY: CEEB, 1980, p. 5. Copyright 1980 by College Entrance Examination Board. Used with permission.

---

freshmen English courses, the scale has been truncated at 60+. [10:7-8; 6.2:5]

Although males have had higher average scores on the SAT-V since 1972 and on the SAT-M since 1967, females have had higher average scores on the TSWE since its introduction in 1975. Between 1975 and 1980, the average score of women on the TSWE declined from 44.3 to 43.0. (See Figure 3.) During the same period, the average score of men declined from 42.2 to 41.7. The average score for both sexes combined has declined from 43.2 to 42.4, a decrease of 0.8 test points. [10:8; 9:7; 8:6; 7:5-6; 6.2:5-6] According to ETS, the 50 test questions are worth approximately 1.0 points each. Thus, the decrease of 0.8 points for both sexes combined during the last five years corresponds to slightly less than one test question.

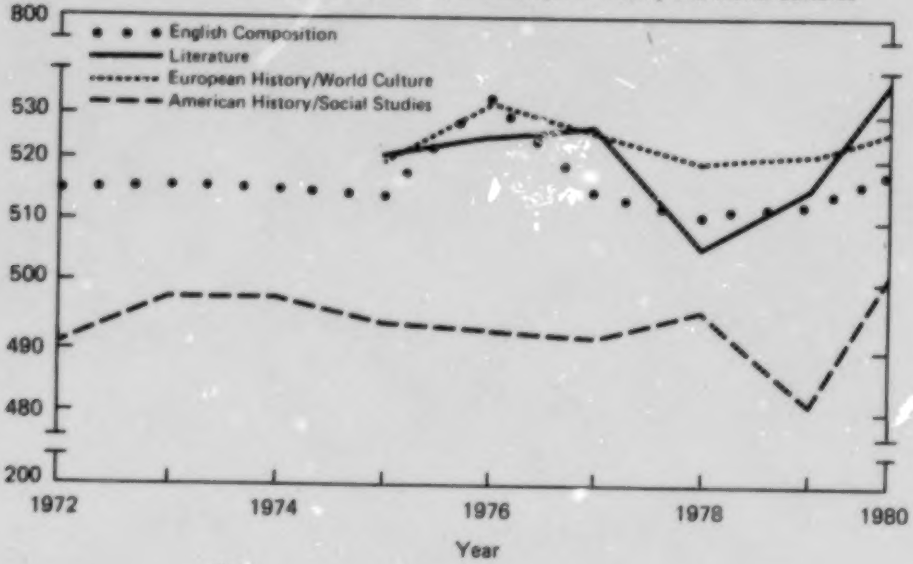
The standard error of the difference for the TSWE administered in May 1978 (Form 3) was 3.7. Therefore, test score differences of less than 7.8 (5.2 x 1.5) should not be considered significant when comparing the scores of two individuals on the TSWE. [6:10]

TABLE 4.--Achievement Test Score Averages, 1972-1980

	1972	1973	1974	1975	1976	1977	1978	1979	1980
Average for All Achievement Tests	526	527	533	531	538	533	531	529	532
American History/ Social Studies	492	498	498	494	493	492	496	480	501
Biology	535	532	545	544	543	543	544	547	551
Chemistry	568	572	581	569	567	574	577	575	573
English Composition	516	517	517	515	532	516	512	514	518
European History/ World Cultures	n/a	n/a	n/a	521	531	526	507	516	539
French	539	544	560	553	553	553	552	554	550
German	n/a	n/a	n/a	547	555	551	553	550	552
Hebrew	n/a	n/a	n/a	577	579	581	589	588	600
Latin	n/a	n/a	n/a	514	524	517	508	524	529
Literature	n/a	n/a	n/a	522	525	526	521	522	524
Mathematics Level I	541	537	545	545	546	547	541	537	536
Mathematics Level II	n/a	n/a	n/a	660	665	666	665	657	653
Physics	n/a	n/a	n/a	601	592	593	591	580	592
Russian	n/a	n/a	n/a	540	559	575	587	613	622
Spanish	530	539	560	544	547	535	544	542	524

SOURCES: Admissions Testing Program of the College Board. *College-Bound Seniors*. New York, NY: CEEB, *1977* (1977), p. 8; *1978* (1978), pp. 13-14; *1979* (1979), pp. 13-14; *1980* (1980), pp. 13-14. Copyright 1977, 1978, 1979, 1980 by College Entrance Examination Board. Used with permission.

FIGURE 4.--Achievement Test Score Averages, 1972-1980: English Composition, Literature, American History/Social Studies, and European History and World Cultures



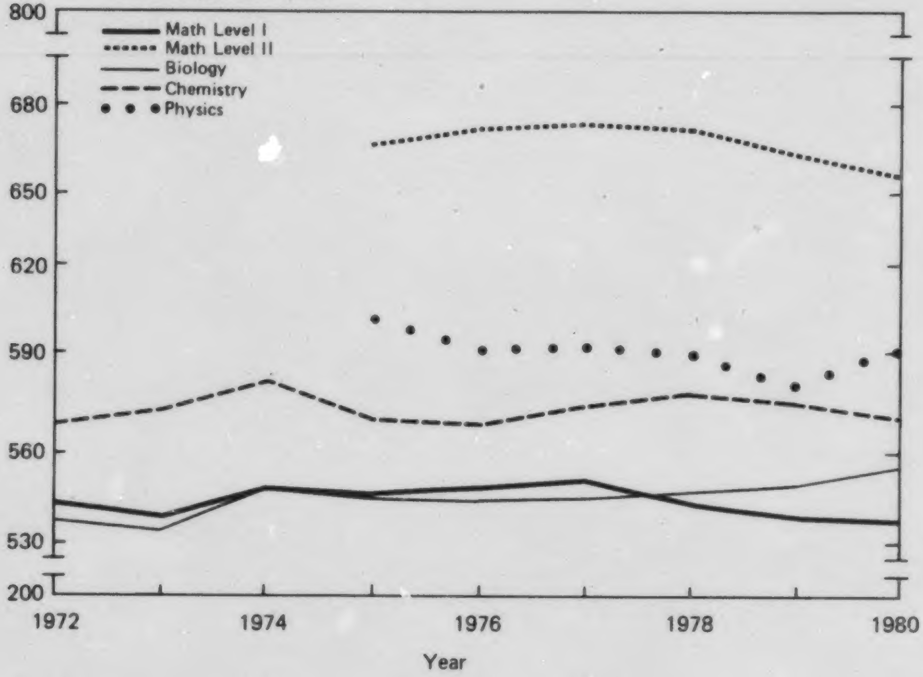
SOURCES: Admissions Testing Program of the College Board. *College-Bound Seniors.* New York, NY: CEEB, *1977* (1977), p. 8; *1978* (1978), pp. 13-14; *1979* (1979), pp. 13-14; *1980* (1980), pp. 13-14. Copyright 1977, 1978, 1979, 1980 by College Entrance Examination Board. Used with permission.

Achievement tests.-- With the exception of the English composition test, which has a 40-minute multiple-choice section and a 20-minute essay on the December edition, the achievement tests are one-hour tests in 15 academic subjects. Some achievement test scores are used in the admissions process while others are used for placement and guidance. [6.1:4] Although more than one million candidates take the SAT annually, fewer than one in five took at least one achievement test in 1980. [6.2:11] Students who take the achievement tests are characteristically of higher ability, according to the College Board. [6.2:6] Unlike the continuous decline in SAT scores, Table 4 shows that the achievement tests present a mixed picture of score gains, declines, and stability.

Figure 4 indicates the trends in the English Composition, Literature, American History/Social Studies, and European History and World Cultures achievement test scores from 1972 to 1980.

- *English Composition:* With the exception of an aberration in 1976, test scores on the English Composition test appear to be stable. The 1980 average of 518 is two points higher than the 1972 average. There are 95 questions on the older editions of the test, 90 on the newer, and 70 when the essay component is included. Each test question is worth approximately six points. A difference of less than 71 points should not be considered significant when comparing the test scores of two students. When the essay component

FIGURE 5.--Achievement Test Score Averages, 1972-1980: Math Level I, Math Level II, Biology, Chemistry, and Physics



SOURCES: Admissions Testing Program of the College Board. *College-Bound Seniors.* New York, NY: CEEB, *1977* (1977), p. 8; *1978* (1978), pp. 13-14; *1979* (1979), pp. 13-14; *1980* (1980), pp. 13-14. Copyright 1977, 1978, 1979, 1980 by College Entrance Examination Board. Used with permission.

is included, test scores of less than 69 to 81 points are not considered significant.

- *Literature:* Literature test scores experienced small fluctuations between 1975 and 1980. The 1980 average of 524 is two points higher than the 1975 average. Each of the 60 questions is worth approximately nine test points on the College Board's 200-to-800 scale. Only when the scores of two students differ by 80 points or more can it be assured that there are real differences in student abilities.
- *American History and Social Studies:* Compared to the 1979 average, the 1980 American History and Social Studies test score increased 21 points to 501, its highest recorded level. Each of the 100 test questions are worth approximately six points. When comparing the scores of students on this test, differences of fewer than 68 points are not considered significant.
- *European History and World Cultures:* In 1980 the European History and World Cultures test score average reached an all time high of 539, up 23 points from the previous year. The point value of each of the 100 test questions is fixed at six points. Differences of less than 66 points in the scores of two students on this test should not be considered significant.

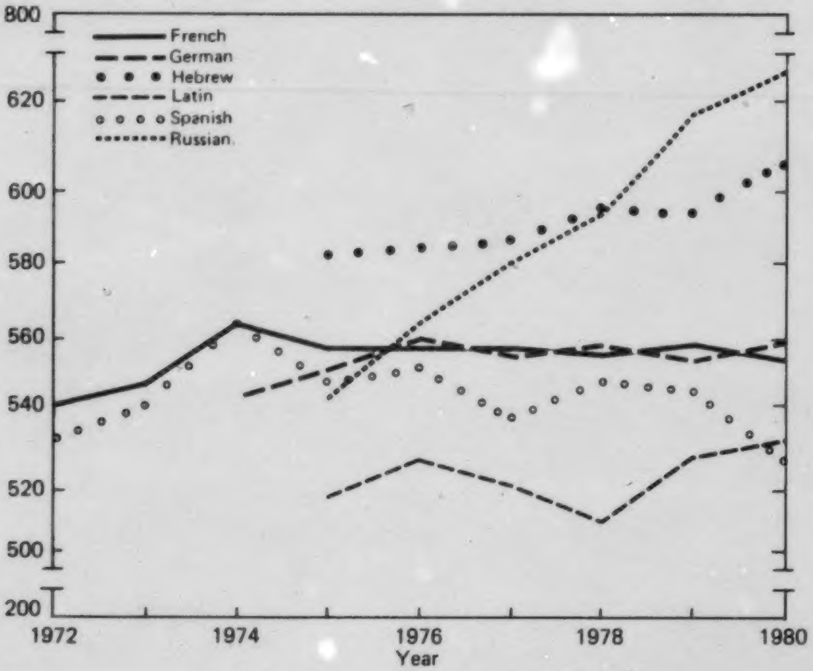
Figure 5 indicates the trends in the Biology, Chemistry, Mathematics Level I, Mathematics Level II, and Physics achievement test scores from 1972 to 1980.

- *Biology:* In 1980, biology test scores averages increased for the third consecutive year. The 1980 average of 551 is the highest level recorded and is 16 points above the 1972 average of 535. The value for the 100 questions is fixed at six points. When comparing the test scores of two students, a difference of more than 63 points is necessary in order to determine that one student has greater ability than another.
- *Chemistry:* Although the chemistry test score average has declined for the second year in a row, the 1980 average of 573 is six points higher than the 1976 average of 567. Each of the 90 questions has a fixed value of six points. Differences in test scores of more than 59 points are necessary to distinguish between the abilities of two students being measured by this test.
- *Mathematics Level I*: Since 1977, test score averages on the Mathematics Level I test have declined 11 points from 547, the all-time high. Viewed from a longer perspective, the 1980 average is five points lower than the 1972 average of 541. Each of the 50 Mathematics Level I questions is fixed at 10 points. A difference of less than 75 points, when comparing two students' scores, is not considered significant.
- *Mathematics Level II:* For the third consecutive year, the Mathematics Level II test score average declined. In 1980 it reached its lowest recorded level, 653, but remains the highest average among all the achievement tests. Point values for each of the 50 questions are fixed at nine points. Only when test score averages exceed 72 points can one be sure that there is a difference in the abilities of two students being tested.
- *Physics:* The 1980 physics test score average of 592 is 12 points higher than the 1979 average, but nine points below the 1975 average of 601. Each of the 75 questions has a value of six points. Test score differences between two students of more than 62 points are needed to be certain that there is a real difference in student abilities.

Figure 6 indicates the trends in the French, German, Hebrew, Latin, Russian, and Spanish achievement test scores from 1972 to 1980.

- *French:* The 1980 average on the French test of 550 is four points below the 1979 average, but is 11 points higher than the 1972 average of 539. Each of the 85 French questions is worth approximately six points. Differences in two students' scores of less than 60 points are not considered significant.
- *German:* Test score averages in German have been relatively stable. Since the 1976 average of 553, test scores have fluctuated five points. The 1980 average is 552. Each of the 80 questions is worth approximately six points. Only when differences in students' test scores exceed 57 points can one be reasonably sure that the higher score indicates higher developed abilities.
- *Hebrew:* With the exception of a decline of one test point in 1979, the Hebrew test score average has climbed consistently in the last six years. The 1980 average of 600 was the highest recorded level, 23 points more than the 1975 average. Each of the 90 questions is worth approximately four points. A difference of more than 42 points in the

FIGURE 6.--Achievement Test Score Averages, 1972-1980: French, German, Hebrew, Latin, Russian, and Spanish



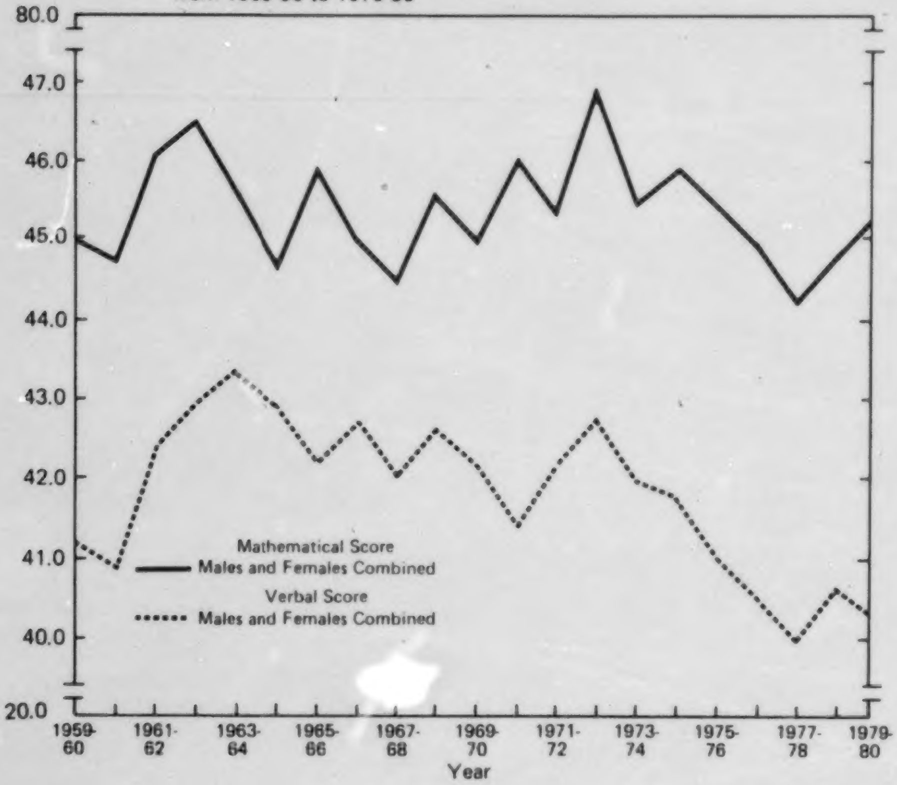
SOURCES: Admissions Testing Program of the College Board. *College-Bound Seniors.* New York, NY: CEEB, *1977* (1977), p. 8; *1978* (1978), pp. 13-14; *1979* (1979), pp. 13-14; *1980* (1980), pp. 13-14. Copyright 1977, 1978, 1979, 1980 by College Entrance Examination Board. Used with permission.

test scores of two students shows that the higher score represents greater abilities on this test.

- *Latin:* In 1980 the Latin test score average reached an all-time high of 529. Each of the 75 questions is worth approximately six points. Differences of less than 77 points in the scores of two students being measured should not be considered significant.
- *Russian:* Since 1975 Russian test score averages have climbed substantially, rising from 540 to 622, an increase of 82 points. Each of the 80 questions is worth approximately six points. Only when test scores differ by more than 60 points can one be sure that there is a real difference in the abilities of two students being tested.
- *Spanish:* Spanish test score averages were erratic during the 1970s. Compared to the 1979 average, test scores declined sharply in 1980. This 18 point decline to 524 was their lowest recorded level. Each of the 85 questions is worth approximately five points. A difference of fewer than 56 points in the scores of two students on this test is not considered significant.

The most popular subjects selected by candidates taking the achievement test series in 1980 were the English composition test (184,714 students), Mathematics Level I test (146,172 students), and the American History and Social Studies test (55,987 students). The College Board reported that the most able candidates

FIGURE 7.--Preliminary Scholastic Aptitude Test/National Merit Scholarship Qualifying Test Means of Mathematical and Verbal Scores for Male and Female Juniors Tested from 1959-60 to 1979-80



SOURCE: College Entrance Examination Board.

selected the Mathematics Level II, physics, Latin, and chemistry tests; the least able candidates took the Spanish, literature, Mathematics Level I, and American history tests. [6.2:6, 13-14]

Overall, the Mathematics Level I and Mathematics Level II test score averages have exhibited a slight downward trend in the late 1970s and into 1980. Test score averages in European history and world cultures, Spanish, and Latin have shown erratic trends. With the exception of a one year aberration, averages on the English composition, American history and social studies, and physics tests appear somewhat stable. Chemistry, literature, French, and German also appear stable. On the other hand, achievement test score averages in biology and Hebrew have shown upward movement in the late 1970s and into 1980, while Russian test scores have climbed dramatically.

Table 5 provides a summary of the standard error of the difference for the 15 achievement tests. The College Board cautions both high school counselors and college admissions officers against "making fine distinctions between scores." Differences in test scores of less than 1.5 times the standard error of the difference are not significant when comparing the scores of two individuals on the same test. [6:11]

Preliminary Scholastic Aptitude Test/National Merit Scholarship Qualifying Test.-- Since 1971 the College Entrance Examination Board and the National Merit Scholarship Corporation (NMSC) have

TABLE 5.--Reliability Characteristics of the College Board's Achievement Tests

Individual ATP Test	Standard Error of the Difference	Differences in ATP Test Scores (Standard Error of the Difference x 1.5)
American History/ Social Studies	45	68
Biology	42	63
Chemistry	39	59
English Composition	47	71
English Composition with Essay	46-54	69-81
European History and World Cultures	44	66
French	40	60
German	38	57
Hebrew	28	42
Latin	51	77
Literature	53	80
Math Level I	50	75
Math Level II	48	72
Physics	41	62
Russian	40	60
Spanish	37	56

SOURCE: Admissions Testing Program of the College Board. *ATP Guide for High Schools and Colleges 1979-81.* New York, NY: CEEB, p. 11. Copyright 1979, 1980 by College Entrance Examination Board. Used with permission.

jointly sponsored the Preliminary Scholastic Aptitude Test/National Merit Scholarship Qualifying Test (PSAT/NMSQT). The PSAT/NMSQT is used to aid guidance counselors in assessing students' competency relevant to college admissions, identify high-scoring students for NMSC scholarships, and prepare juniors for taking the SAT. The PSAT/NMSQT is a multiple-choice test constructed from retired forms of the SAT. [19:45; 56:6-8] It consists of 65 verbal questions and 50 mathematical questions, with each verbal question worth approximately 0.9 points and each mathematical question approximately 1.0 points. Verbal and mathematical scores for the PSAT/NMSQT are reported on a 20-to-80 scale. Like the Test of Standard Written English, the PSAT/NMSQT scores correspond to the SAT scale of 200 to 800. Juniors and seniors can estimate their probable SAT scores by adding a zero to the PSAT/NMSQT score. For example, a PSAT verbal score of 45 corresponds to a SAT verbal score of 450. [55:7]

Table 6 indicates the numbers, means, and standard deviations of verbal scores for juniors tested from 1959-60 to 1979-80. Between 1975-76 and 1976-77 the mean PSAT-V scores of males and females combined declined 0.5 points, from 41.0 to 40.5. In 1977-78 the verbal score declined 0.6 points to 39.9, the lowest score ever recorded. The verbal score rebounded by 0.7 points, to 40.6, in 1978-79, but fell 0.4 points, to 40.2, in 1979-80. Over the recent four-year testing period, the PSAT-V score has fluctuated by 1.1 points, compared to a fluctuation of 3.4 points over the 20-year history of the test's administration. The most recent PSAT-V score of males and females combined in 1979-80 is 1.0 points below the score of 41.2 in 1959-60 and 1.2 points below the score of 41.4 in 1970-71

TABLE 6.--Verbal Scores for Juniors on the Preliminary Scholastic Aptitude Test/National Merit Scholarship Qualifying Test 1959-60 to 1979-80*

Numbers, Means, and Standard Deviations of Verbal Scores for Junior Students Tested from 1959-60 to 1979-80

	Male			Female			Total		
Year	Number	Mean	SD	Number	Mean	SD	Number	Mean	SD
1959-60	186,791	41.0	11.2	191,365	41.3	10.8	378,156	41.2	11.0
1960-61	229,781	40.9	10.7	232,008	41.0	10.4	461,789	40.9	10.6
1961-62	253,001	42.1	10.7	254,839	42.6	10.6	507,839	42.3	10.7
1962-63	317,055	42.8	10.7	330,624	42.9	10.5	647,709	42.9	10.5
1963-64	379,282	43.1	10.8	395,793	43.4	10.6	775,075	43.3	10.7
1964-65	402,465	42.7	10.6	421,496	43.0	10.4	823,961	42.9	10.5
1965-66	447,808	42.5	11.0	477,685	42.0	10.9	926,035	42.2	10.9
1966-67	461,994	42.7	11.0	496,478	42.7	10.7	958,974	42.7	10.8
1967-68	493,831	42.0	10.8	538,421	42.1	10.7	1,032,945	42.0	10.8
1968-69	508,888	42.2	11.5	555,304	42.9	11.6	1,064,986	42.6	11.6
1969-70	507,570	42.2	10.8	559,721	42.2	10.8	1,067,834	42.2	10.8
1970-71	492,990	41.4	11.0	553,583	41.4	10.7	1,047,135	41.4	10.8
1971-72	512,622	42.1	10.6	576,393	42.4	10.5	1,090,868	42.2	10.6
1972-73	491,385	42.6	11.1	564,975	42.7	10.9	1,056,752	42.7	11.0
1973-74	492,683	42.0	10.8	574,424	41.7	10.7	1,067,601	41.8	10.8
1974-75	503,959	42.2	10.7	589,840	41.1	10.4	1,094,328	41.6	10.5
1975-76	508,834	41.8	11.1	602,818	40.4	11.0	1,112,483	41.0	11.1
1976-77	510,543	40.9	10.4	609,174	40.1	10.3	1,120,542	40.5	10.4
1977-78	521,802	40.6	10.6	629,870	39.3	10.6	1,152,494	39.9	10.6
1978-79	513,515	41.2	10.4	622,212	40.1	10.3	1,136,586	40.6	10.3
1979-80	510,115	40.7	10.3	621,493	39.9	10.4	1,132,370	40.2	10.4

*Preliminary Scholastic Aptitude Test, 1959-60 through 1970-71; Preliminary Scholastic Aptitude Test/National Merit Scholarship Qualifying Test from 1971-72 to 1979-80.

SOURCE: College Entrance Examination Board.

Table 7 shows the PSAT/NMSQT numbers, means, and standard deviations of mathematical scores for juniors tested from 1959-60 to 1979-80. The mean PSAT-M score for males and females combined was 45.5 in 1975-76. It declined 0.5 points to 45.0 in 1976-77 and 0.8 points to 44.2 in 1977-78. During the past two years, however, the PSAT-M has risen by 1.1 points to 45.3. Since 1975-76 the mathematical score has fluctuated by 1.3 points, as compared to 2.7 points over the entire history of the PSAT's administration. The 1979-80 PSAT-M mean of 45.3 is 0.3 points higher than the 1959-60 mean of 45.0, but 0.8 points lower than the 1970-71 mean of 46.1.

Presented in Figure 7 is a graphical representation of the yearly fluctuations of the PSAT/NMSQT from 1959-60 to 1979-80 for males and females combined. Unlike the continuing decline in SAT scores, the PSAT scores for juniors shown in this graph have fallen much less and also less uniformly than scores on the SAT. In 1970-71, PSAT/NMSQT verbal scores for males and females combined were 41.4. They declined by 1.2 points, to 40.2, in 1979-80. This decline of 1.2 is equivalent to one and one-third fewer test questions answered correctly by test takers during the 1970s. Similarly, the 1970-71 PSAT/NMSQT math scores for males and females

TABLE 7.--Mathematical Scores for Juniors on the Preliminary Scholastic Aptitude Test/National Merit Scholarship Qualifying Test 1959-60 to 1979-80*

Numbers, Means, and Standard Deviations of Mathematical Scores for Junior Students Tested from 1959-60 to 1979-80

Year	Male			Female			Total		
	Number	Mean	SD	Number	Mean	SD	Number	Mean	SD
1959-60	186,791	47.6	10.2	191,365	42.4	8.6	378,156	45.0	9.8
1960-61	229,755	47.5	10.2	231,956	42.3	9.2	461,711	44.8	10.1
1961-62	252,995	48.6	10.4	254,820	43.6	9.4	507,804	46.1	10.2
1962-63	317,069	48.7	10.6	330,599	44.3	9.7	647,668	46.5	10.4
1963-64	379,266	47.6	10.5	395,793	43.5	9.3	775,075	45.6	10.1
1964-65	402,465	46.9	10.9	421,495	42.6	9.8	823,961	44.7	10.6
1965-66	447,803	48.3	11.1	477,678	43.8	10.2	926,023	45.9	10.9
1966-67	461,936	47.1	11.5	496,455	43.1	10.7	958,943	45.0	11.3
1967-68	493,829	46.4	10.8	538,418	42.7	9.7	1,032,940	44.5	10.4
1968-69	508,858	47.4	11.5	555,260	44.0	10.5	1,064,912	45.6	11.1
1969-70	507,557	47.0	11.4	559,714	43.2	10.6	1,067,814	45.0	11.1
1970-71	492,982	48.4	11.5	553,568	44.1	10.6	1,047,112	46.1	11.2
1971-72	512,622	47.3	11.6	576,393	43.4	10.5	1,090,863	45.2	11.2
1972-73	491,294	49.1	11.4	564,864	45.0	10.4	1,056,550	46.9	11.1
1973-74	492,573	47.8	11.8	574,256	43.5	10.9	1,067,323	45.5	11.5
1974-75	503,790	48.2	11.2	589,620	43.9	9.9	1,093,938	45.9	10.7
1975-76	508,700	47.9	11.4	602,631	43.5	10.4	1,112,159	45.5	11.1
1976-77	510,435	47.3	11.5	609,036	43.0	10.4	1,120,298	45.0	11.1
1977-78	521,631	46.5	10.9	629,669	42.2	9.8	1,152,122	44.2	10.5
1978-79	513,382	47.2	11.5	622,041	42.9	10.2	1,136,281	44.8	11.0
1979-80	509,981	47.5	11.3	621,338	43.5	10.2	1,132,080	45.3	10.9

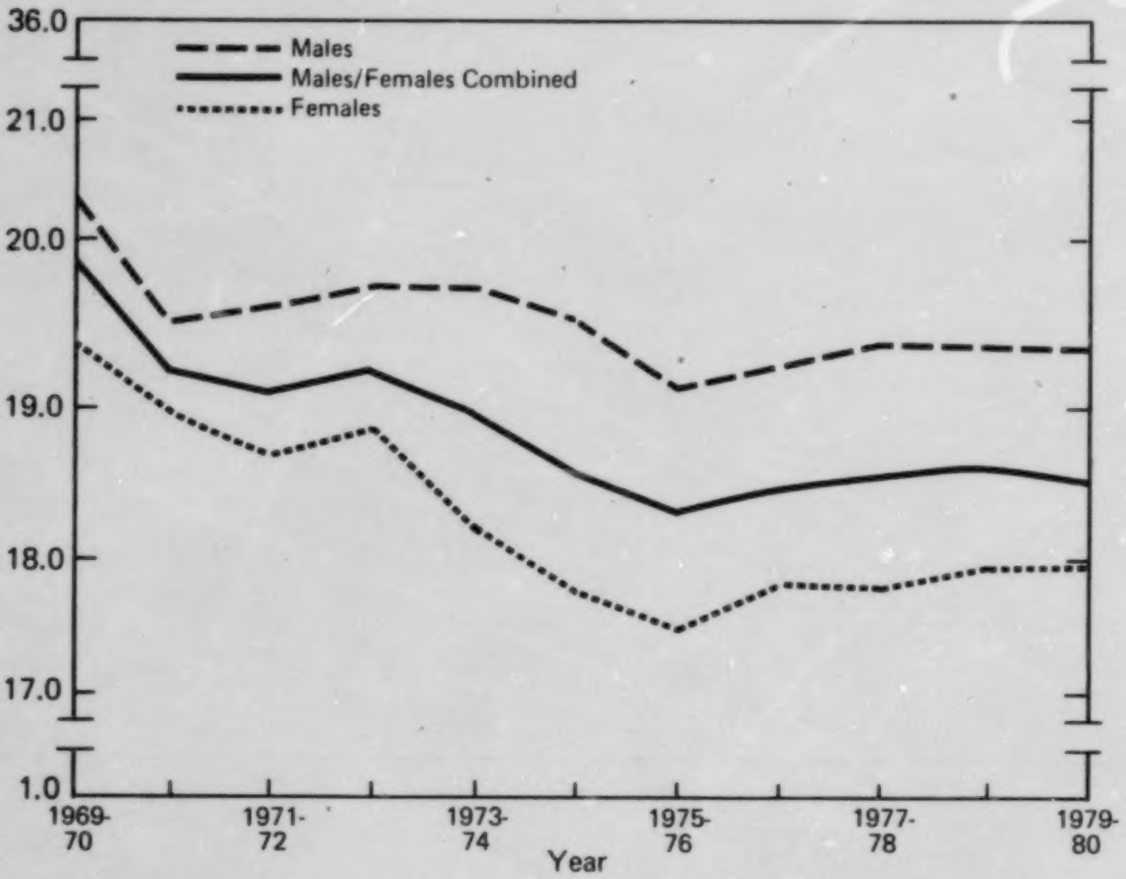
*Preliminary Scholastic Aptitude Test, 1959-60 through 1970-71; Preliminary Scholastic Aptitude Test/National Merit Scholarship Qualifying Test from 1971-72 to 1979-80.

SOURCE: College Entrance Examination Board.

combined were 46.1, compared to 45.3 in 1979-80. The decline of 0.8 points is equal to approximately four-fifths of a test question. Thus, the total decline in PSAT/NMSQT scores during the 1970s was slightly more than two test questions out of a total of 115.

American College Testing Program.-- Since its founding in 1959, the ACT has attempted to assess the general educational development of high school students and their ability to perform college-level work. [16:3-4] More than 900,000 students are tested annually in the United States, Canada, and abroad. Some of the major purposes of the ACT Assessment Program are (1) to provide estimates of students' academic and out-of-class abilities; (2) to give admissions and guidance services to students and admissions personnel during the transition to post-secondary education; (3) to furnish colleges with admissions/enrollment data; (4) to supply pre-college counseling; and (5) to assist in placing freshmen in appropriate sections of introductory college courses in English, mathematics, social studies, and natural sciences. [3:4; 18:2]

FIGURE 8.--Mean Scores from Test Section of ACT Assessment Program: Males, Females, and Males/Females Combined, 1969-70 to 1979-80*



*Mean scores compiled from data based on 10 percent sample of students who took the ACT Assessment's four subject area tests on national test dates each year. Only first four test dates each year were used from 1969-70 through 1975-76. Since then, all five dates each year are used.

SOURCES: The American College Testing Program. "ACT Assessment Test Scores-3rd Year in Row for Increase," *Activity*, 18 (May 1980), p. 4, and letter from ACT (October 1980). Used with permission.

One component of the ACT Assessment Program is a battery of four academic tests: English Usage, Mathematics Usage, Social Studies Reading, and Natural Sciences Reading. Test items require students to demonstrate reasoning abilities and problem solving skills, as well as knowledge in the four subject areas. Specifically, the English Usage Test is a 40-minute, 75-item test designed to measure students' understanding of the conventions of standard written English and the elements of expository writing: logic and organization, sentence structure, diction, style, grammar, and punctuation. The Mathematics Usage Test is a 50-minute, 40-item test designed to measure mathematical reasoning ability. It emphasizes the solution of quantitative problems. The Social Studies Reading Test is a 35-minute, 52-item test designed to measure knowledge and problem solving skills required in social studies. Some test items are based on reading passages and require reading comprehension and the ability to make inferences and deductions; examine interrelationships; recognize a writer's bias, style, and mode of reasoning; and draw conclusions. Other items test the ability to apply previously acquired knowledge to new situations. The Natural Sciences Reading Test is a 35-minute, 52-item test designed to measure knowledge and the ability to read, analyze, and evaluate natural science material. The passages cover a variety of scientific topics found in the high school curriculum. [18:3]

Test scaled scores on the ACT range from a low of 1 to a high of 36, but the minimum Composite score is 1 and the maximum Composite score is 35. The standard error of measurement for each of the ACT tests is about 2 and for the Composite score, about 1. According to the ACT Assessment Program, an English Usage Test score of 18 should fall within a range of 16 to 20, and a Composite score of 19 should fall within a range of 18 to 20. [18:4] A Composite score below 14 indicates that a student has had "a restricted educational development background." Scores between 14 and 19 are considered low average; between 19 and 24, high average; and above 24, superior. [18:12] Since there is no penalty for guessing on the test, it is to the student's advantage to answer every question. [18:5]

Table 8 and Figure 8 show the mean scores from the test section of the ACT Assessment Program from 1969-70 to 1979-80. The data are based on a 10 percent sample of students from each year's national test dates, rather than on all students tested in a given year. This sample is considered to reflect better the college-bound students who participate in the ACT Assessment Program each year. [3:4]

The ACT Assessment Program experienced a decline in test scores similar to the decline in SAT test scores in the early and mid-1970s. But between 1975-76 and 1978-79, the ACT Composite scores of males and females combined steadily increased, unlike SAT scores, rising from 18.3 to 18.6. This increase over a three-year period was noteworthy, according to the ACT, because only once since 1969-70 had there been an increase in the mean Composite score, which was an increase of one-tenth of a standard score point in 1972-73. [3:4] Most recently, the 1980 Composite score of males and females combined declined by one-tenth of a standard score point to 18.5, the same level recorded in 1977-78. According to Richard L. Ferguson, senior vice president of ACT Programs and Services, "small variations from one year to another are not at all uncommon. Only when they yield consistently lower or higher mean scores over an extended period of time are we able to determine the existence of a trend." [3:4] A discussion of ACT subtest score movement in the 1970s follows.

- The Composite score of *males* increased from a low of 19.1 in 1975-76 to 19.3 in 1977-78. It has remained unchanged at 19.3. The Composite score of *females* rose from 17.6 in 1975-76 to 17.8 in 1976-77. It remained unchanged the following year, rising to 17.9 in 1978-79, where it remained unchanged the following year.
- The *English* subtest score of males increased consistently from a low of 16.8 in 1975-76 to 17.4 in 1977-78. It remained unchanged at 17.4 in 1978-79 before declining to 17.3 in 1979-80. The English subtest score of females increased steadily from a low of 18.0 in 1975-76 to

TABLE 8.--Mean Scores from Test Section of ACT Assessment Programs

Year	English	Math	Social Studies	Natural Sciences	Composite
		Males/Females Combined			
1969-70	18.5	20.0	19.7	20.8	19.9
1970-71	18.0	19.1	18.7	20.5	19.2
1971-72	17.9	18.8	18.6	20.6	19.1
1972-73	18.1	19.1	18.3	20.8	19.2
1973-74	17.9	18.3	18.1	20.8	18.9
1974-75	17.7	17.6	17.4	21.1	18.6
1975-76	17.5	17.5	17.0	20.8	18.3
1976-77	17.7	17.4	17.3	20.9	18.4
1977-78	17.9	17.5	17.1	20.9	18.5
1978-79	17.9	17.5	17.2	21.1	18.6
1979-80	17.9	17.4	17.2	21.1	18.5
		Males			
1969-70	17.6	21.1	20.3	21.6	20.3
1970-71	17.1	20.2	19.0	21.3	19.5
1971-72	17.0	20.1	19.1	21.6	19.6
1972-73	17.3	20.2	19.0	21.7	19.7
1973-74	17.1	19.7	19.1	22.2	19.7
1974-75	17.1	19.3	18.7	22.4	19.5
1975-76	16.8	19.2	17.9	22.0	19.1
1976-77	17.0	18.9	18.2	22.3	19.2
1977-78	17.4	19.1	18.0	22.3	19.3
1978-79	17.4	19.1	18.1	22.3	19.3
1979-80	17.3	18.9	18.2	22.4	19.3
		Females			
1969-70	19.4	18.8	19.0	20.0	19.4
1970-71	19.0	18.0	18.4	19.7	18.9
1971-72	18.7	17.7	18.2	19.6	18.7
1972-73	18.9	18.0	17.7	19.9	18.8
1973-74	18.6	17.1	17.3	19.6	18.2
1974-75	18.3	16.2	16.4	20.0	17.8
1975-76	18.0	16.0	16.2	19.7	17.6
1976-77	18.2	16.1	16.5	19.6	17.8
1977-78	18.3	16.2	16.4	19.8	17.8
1978-79	18.4	16.2	16.4	20.2	17.9
1979-80	18.3	16.2	16.4	20.0	17.9

Mean scores compiled from data based on 10 percent sample of students who took the ACT Assessment's four subject-area tests on national test dates each year. Only first four test dates each year were used from 1969-70 through 1975-76. Since then, all five dates each year are used.

SOURCE: The American College Testing Program. "ACT Assessment Test Scores--3rd Year in Row for Increase," *Activity*, 18 (May 1980), p. 4, and letter from ACT (October 1980). Reprinted with permission.

18.4 in 1978-79. It declined one-tenth of a standard score point to 18.3 in 1979-80.

- The *mathematics* subtest score of males declined three-tenths of a standard score point from 19.2 in 1975-76 to 18.9 in 1976-77. It increased to 19.1 in 1977-78 and remained unchanged the following year. In 1979-80 the subtest score of males declined two-tenths of a standard score point to 18.9, the same level recorded in 1976-77. The mathematics subtest score of females increased consistently from 16.0 in 1975-76 to 16.2 in 1977-78. It has remained unchanged.
- The *social studies* subtest score of males increased three-tenths of a standard score point from 17.9 in 1975-76 to 18.2 in 1976-77. It declined to 18.0 in 1977-78 and increased to 18.1 the next year. It rose to 18.2 in 1979-80. The social studies subtest score of females increased three-tenths of a standard score point from 16.2 in 1975-76 to 16.5 in 1976-77. It declined to 16.4 in 1977-78, where it has remained unchanged.
- The *natural sciences* subtest score of males increased from 22.0 in 1975-76 to 22.3 in 1976-77, remaining unchanged through 1978-79. It increased one-tenth of a standard score point to 22.4 in 1979-80. The natural sciences subtest score of females declined one-tenth of a standard score point from 19.7 in 1975-76 to 19.6 in 1976-77. It increased to 19.8 in 1977-78 and then by four-tenths of a standard score point, to 20.2, in 1978-79, its highest recorded level in the 1970s. In 1979-80 it declined by two-tenths of a standard score point to 20.0.

The most recent ACT national class profile norms were based on a 10 percent sample of ACT-tested students enrolled at 1,103 colleges in 1979-80. The norms, however, overrepresent ACT-tested college-bound students in the Midwest, South, Rocky Mountains, and Plains states and underrepresent college-bound students in the Northeast and Middle Atlantic states. Likewise, private colleges and universities are underrepresented in the sample. [15:1]

- Of the 46,606 students in the sample, 21,333 were males and 25,273 were females. [15:1]
- The typical student in the 1979-80 sample had an ACT Composite score of 18.9 and a high school average of 3.0, compared to the national averages of 18.7 and 3.0, respectively, for 1978-79 entering freshmen. [15:1]
- Students were primarily interested in the field of business-commerce for their planned educational major and first vocational choice. [15:1]
- More students aspired to a bachelor's degree than to a graduate or professional degree, 44 percent to 39 percent. [15:1]
- Regarding the racial composition of the sample, 75 percent identified themselves as Caucasian American; seven percent, Afro-American/Black; two percent, Mexican American/Chicano; one percent, American/Alaskan Native; one percent, Oriental/Pacific American; less than one percent, Puerto Rican/Hispanic; two percent, other; four percent preferred not to respond; and eight percent were not given. [15:1]
- Students expressed a need for help in the following areas: 44 percent with educational and vocational plans; 40 percent with mathematics; 37 percent with study skills; 32 percent with reading; 30 percent with personal counseling; and 27 percent with writing. [15:1]
- A majority of the students, 55 percent, expressed a need for assistance in finding a school-year job. [15:1]

WHY THE TRENDS IN TEST SCORES

With college admission test scores declining during the 1970s, many educators searched for possible causes of this decline. One of the most interested parties in discovering why test scores were decreasing was the College Board. This section discusses the background and results of a major study supported by the College Board, *On Further Examination*; criticisms of this report; and theories advanced since the publication of *On Further Examination* on possible causes for declining test scores.

On Further Examination

In October 1975, S. P. Marland, Jr., then president of the College Entrance Examination Board, appointed a 21-member "blue-ribbon" Advisory Panel on the SAT score decline. The Advisory Panel was empowered to investigate the complex and interrelated issues pertinent to the unexplained decline in SAT scores, examine the psychometric integrity of the tests, and identify areas of additional research necessary for dealing effectively with the test score decline. Chaired by Willard W. Wirtz, former U.S. Secretary of Labor, the Advisory Panel was given freedom to deliberate issues as it chose and to report its findings in the public interest. [297:iii] After a two-year investigation at a cost of $600,000, the Advisory Panel's report, *On Further Examination* and new research, were released in 1977. [107:61] Since the issues in the report were of national interest, the Panel emphasized that any generalizations regarding the SAT statistics should be carefully qualified, and "should not be extended to cover the situation of American youth as a whole or the overall effectiveness of the learning process." [297:5]

The Advisory Panel concluded that there were actually two separate SAT score declines characterized by different causal factors. The first decline, between 1963 and 1970, was due primarily to changes in the SAT-taking population. Compared to the past, the SAT was measuring a broader cross section of American youth which included larger proportions of characteristically lower scoring groups of students from disadvantaged socio-economic backgrounds, ethnic minorities, and women. The second decline in SAT scores, between 1970 and 1975, was due to pervasive changes affecting both higher- and lower-scoring groups. [297:13, 45-46] Specifics of the first decline in SAT scores related to compositional changes in the SAT-taking population follow:

- SAT score averages began to decline with the post-World War II wave of students. Between 1960 and 1970 the number of SAT-takers tripled, and the number of 18-year-olds increased by one million between 1964 and 1965. [297:13]
- Twenty-five years ago, about half of all students were remaining in school through the twelfth grade. By 1970 the percentage

had grown to three-fourths. The increase in young people remaining in school resulted in a "lowering of the *averages* on examinations taken previously by more select groups." [emphasis in the original] [297:13-14, 24]

- Students from families with lower-than-average incomes, under $6,000 in 1977, averaged about 100 points lower on both the verbal and mathematical sections of the SAT compared to students from high-income families ($18,000 and over). [297:15]
- Few ethnic minority groups took the SAT in the early 1960s. Black SAT-takers in 1963 comprised between one and two percent of the SAT population. By 1972 this population had risen to 8.7 percent. Blacks averaged approximately 100 points below the national average on the verbal test and about 115 points below on the mathematical section of the SAT. [297:15]
- The percentage of women taking the SAT rose from 42.7 percent in 1960 to 51.1 percent in 1977. Women's lower scores on the mathematical section of the test, the panel observed, reflected the "traditional sex stereotyping of career opportunities and expectations." [297:16, 19]
- The number of SAT-takers applying to prestigious and selective four-year colleges and universities remained constant until 1967. Thereafter, an increasing *percentage* of test takers began to apply to less selective colleges, colleges with open admission policies, two-year colleges, and schools with a technical or vocational interest.
- Between 1961 and 1974 the percentage of students who requested to have their scores sent to selective colleges dropped from 13.2 percent to 5.6 percent. Similarly, the number of requests to have scores sent to "research" universities declined from 10.5 percent to 6.4 percent. [297:17-18]
- There was also a decrease in the percentage of students repeating the SAT. This had a small effect on test score averages since SAT-repeaters typically averaged 15 to 30 points higher the second time. [297:18]

The Advisory Panel cautioned, however, that identifying the increased numbers of blacks, women, and poorer youths who took the SAT as causes for the test score decline would risk "irresponsible headlines." [297:17] The increased number of blacks would account for a decline of about five points or less over the entire 14-year period. "Despite statutory guarantees of equal opportunity," the Panel said, "the society has not yet developed either the educational means or the mores that will bring children with different racial roots to a parity of aptitude, as the SAT and other tests measure it, by the time they reach the twelfth grade." [297:16] Increased participation of women, they felt, would account for about a four or five point decline in the mathematical average, but no decline in verbal averages. Altogether, the changes in the SAT-taking population, especially "in terms of the test takers coming from higher- and lower-scoring groups and in terms of their plans for going on to college," accounted for an estimated two-thirds to three-quarters of the SAT score decline. [297:17-18]

The second decline in SAT scores, between 1970 and 1975, was related to events in the schools and society at large. The Panel concluded that there was no one cause for the SAT decline but rather a "virtually seamless web of causal connections." [297:25, 48] Among the school and societal factors the Panel cited as contributing to the decline in SAT scores were the following:

- Although well-intentioned, many *curriculum changes* had reduced the continuity of study in major fields with consequent effect on verbal development. "A significant dispersal of learning activities" had occurred in the schools, that is, a reduction in the number of courses required of all students and the addition of elective courses. [297:46] Research cited by the Panel also indicated this trend. For example, Harnischfeger and Wiley, using nationwide data, reported an 11 percent drop in English enrollments in grades 7 through 12 and a 50 percent drop in advanced English between 1971 and 1973. A survey of elective courses in Massachusetts conducted by the Massachusetts Department of Education between 1971 and 1976 found more than a 50 percent increase in English/language arts courses in 43 high schools, but no significant correlation was found between the number of electives added and the students' SAT scores. California reported a 19 percent drop in enrollments in basic English courses between 1971-72 and 1974-75, while enrollments "plummeted" 77 percent in English composition classes. [297:25-26] The Advisory Panel, however, did not condemn elective courses since it felt that many of them contributed to student interest and motivation with no "negative effect on basic learning." [297:46] The relationship between SAT scores and the curriculum led the Panel to observe that "less thoughtful and critical reading is now being demanded and done," and that "careful writing has apparently about gone out of style." [297:27]
- Student *absenteeism* increased considerably with absence rates of 15 percent or more common in some schools. The Panel observed that absenteeism hurt both the absentee and the class, since the teacher often had to repeat a lesson to bring everyone up-to-date. [297:29]
- *Grade inflation* indicated declining educational standards. The Panel cited an American College Testing Program study which reported a 25 percent increase in the proportion of "A" and "B" grades between 1962-65 and 1974-75. Students increasingly accepted the idea that promotion from one grade level to another was "an entitlement rather than something to be earned, or denied." [297:29]
- An analysis of reading, history, and other *textbooks* showed more space devoted to pictures; wider margins; and shorter words, sentences and paragraphs. Since numerous exercises in textbooks had an "objective answer" format, requiring underlining, circling, and filling in blanks, many students found less reason to write. [297:31]
- The Panel stated that the extensive time students spent watching *television*, between 10,000 and 15,000 hours by age 16, detracted from homework and time spent developing skills measured by college entrance examinations. They felt that TV raised children's expectation levels, which by comparison, made what they encountered in the classroom appear rather bland. The Panel noted that although television had become a "surrograte parent, substitute teacher" it still had the potential "of becoming learning's most fertile grove." [297:35-37, 47]
- Changes in the *role of the family*, especially the increasing number of children in single-parent families, may have had a negative impact on students' college entrance examinations. [297:47]
- The Panel said the 17-year-old of 1977 was not the same in mind or body as the

17-year-old of 1947. *Changing life styles*, values, earlier physical maturation, higher mobility, drugs, and contraceptives had coincided with increasing problems of discipline and absenteeism among high school students. However, these factors could not be translated into SAT test score declines. [297:35]

- Whether or not the *disruption of national life* during the 1967-75 period (the Vietnam War, political assassinations, burning cities, corrupt national leadership) affected the motivation of test takers preparing for their college entrance examinations remained unclear. The Panel concluded that the period covered by the second decline in SAT scores was a difficult one in which to grow up. [297:37, 43]
- There was an apparent dimunition in students' *learning motivation*. The Panel noted that the "curve of the SAT scores has followed very closely the curve of the entire nation's spirits, and self-esteem and sense of purpose." [297:48]

In summary, the Panel cautioned that the SAT score decline was a complex subject filled with nuances, qualifications, and doubts. The SAT is a "limited instrument" and should not be viewed as "the sole thermometer for measuring the health of schools, family, and student." [297:40] While high school grades still remained the best single predictor of college performance, the psychometric integrity of the SAT was affirmed as a valid instrument in determining how well students will perform during their first year in college. Furthermore, the Panel noted that the predictive validity of the verbal and mathematical tests had actually increased between 1970 and 1974. [279:9] Problems related to declining test scores were not isolated school problems, but rather society's problems. [297:46] While it appeared that educational standards had diminished, the Advisory Panel wondered why the SAT score decline actually had not been larger. [297:45] The Panel's report concluded on an optimistic note: "We find nothing in the record we have reviewed to discourage the conviction that learning in America can be made all that is hoped for it." [297:48]

Criticisms of On Further Examination

The Advisory Panel's report, *On Further Examination*, found no one cause and no single pattern of causes for the decline in SAT scores. Much of the report blamed society as a whole for declines in educational standards: changes in the role of the family, television, lack of motivation, and a "decade of distraction." [297:37, 48] Since there was little quantitative evidence available about the precise influences of those factors on test scores, many of the Panel's findings were expressed in general terms as opinions and reflections. [193:15-16] Critics faulted the Panel for using "circumstantial evidence" to validate a variety of possible causes for declining admissions test scores. [245:290] Among the criticisms of the Advisory Panel's report were the following:

1. It was merely a "policy paper" which attempted to influence the public schools in the directions suggested by the College Board and the Educational Testing Service. [144:7]
2. Although the Panel emphasized that the SAT's validity as a predictor of success in first-year college performance had actually increased, the SAT still remained "second best in a relatively poor string of predictors." [144:5]
3. Teachers had known first-hand for some time many of the Panel's conclusions regarding changes and conditions in the schools. [239:6]
4. In light of the drastic changes in society which the report documented,

serious questions were raised about the capability of a standarized multiple-choice test designed in 1941 to assess students in the 1960s and 1970s, asserted John Ryor, former president of the National Education Association (NEA). [239:8]

5. Identification of high school electives as a possible contributing factor in the SAT score decline and the Panel's defense of flexibility and diversification were paradoxical points of view. [252:86] While the Panel did not condemn electives, its conclusions effectively did just that. [144:6]
6. School boards that attempt to evaluate their curriculums based on SAT scores were, in effect, making the SAT a "basis for a national standard." [239:8]
7. Instead of attacking the misuse of test score data, the Advisory Panel engaged in "unsubstantiated opinion" about the state of literacy achievement, although SAT scores do not provide adequate data for such judgments. [144:5-6]
8. Arthur N. Applebee of the National Council of Teachers of English observed that English teachers appeared to have been singled out as responsible for the score decline. Applebee said the critical reading and careful writing and thinking that the Panel recommended should be part of the entire curriculum, not just English classes. [20:2]
9. In an editorial in *The English Journal*, Stephen Judy criticized the Panel's "firmest" conclusion that less thoughtful and critical reading was being done and demanded, and that careful writing had gone out of style. Judy queried whether the Panel had actually collected and compared samples of writing to assess the changes and whether the Panel had measured how much "thoughtful and critical reading" is being done. [144:6-7]
10. Despite the Panel's warning to schools about becoming more rigid, restrictive, and uniform, the report encouraged this kind of activity and prompted the cry for "back-to-basics." [239:[illegible]]
11. The Panel's perceived message to the public and politicians was that schools "deserted the basics" and replaced them with innovations, electives, and entertainment. This interpretation, according to one educational observer, has resulted in minimum competency laws in many states. [245:292]
12. The NEA asserted that the Advisory Panel's report raised significant questions about unchanging standards, test validity, cultural bias, and instructional changes--questions that warranted still further examination. [193:18-19]

Recent Theories for Declining Test Scores

Since the release of the Wirtz panel's report, *On Further Examination*, in August 1977, SAT test scores have continued to decline, but not as steeply as in the past. The continuing decline has prompted further speculation about the possible factors influencing test scores. Some educators maintain that there is no way of knowing whether or not there has been a decline in students' academic achievement because of the margin of error in the tests. [60:13] Furthermore, they point to "then and now" studies which indicate that students' reading achievement is about the same today as in the past and in fact compares favorably with students' reading achievement in other countries. [86:3] Other educators claim that the media has confused the SAT score decline with the overall effectiveness of

American education. They believe that a series of misinterpretations has encouraged the public to believe that schools are graduating increasing numbers of illiterates. In addition, some say that this confusion has resulted in a return to teaching the basics and minimum competency testing, without really knowing if this method of instruction is necessary. [87:528] Still other educators believe that these political responses to educational problems tend to focus more attention on testing than on the instructional process and the improvement of verbal and mathematical abilities. [241:2] In any case, the public has demanded a quick cure for what it perceives to be a decline in public education. While the exact cause or causes of declining SAT scores remain speculative, according to some educational analysts, it is far "easier to damage the educational system than it is to improve it." [272:2] Highlights of some recent theories and discussions regarding declining test scores follow.

A March 1978 report of the National Academy of Education Committee on Testing and Basic Skills identified four school factors that had contributed to declines in SAT scores and writing skills: (1) a proliferation of nonintellectually rigorous courses less demanding than the traditional regimen of English, math, science, and social studies, accompanied by a decline in the intellectual standards of formerly rigorous, but still required courses; (2) confusion about the appropriate roles of teachers and "often contradictory modes of appropriate pedagogic behavior"; (3) a decrease in the amount of time spent "on task," with some "good" classes devoting only 60 or 70 percent of their instructional day to learning; and (4) less opportunity for intensive study by academically talented students at the secondary level, which contributed to the decline in the number of high-scoring students on the SAT. The National Academy of Education panel emphasized that the effective use of school time spent "on task" in whatever the curriculum called for would be a "significant reform in American schools." [190:1-2]

Vito Perrone, dean of the Center for Teaching and Learning at the University of North Dakota, criticized the National Academy of Education's report as laden with overgeneralizations and opinions, with little substantiated data to support the report's recommendations. Perrone said there was no empirical evidence that a proliferation of nonintellectually rigorous courses, confusion about the appropriate role of teachers, slackening of "on task" attention, and a dismantling of opportunity for intensive study leads to declines in writing skills and SAT scores. Perrone reported that data from the International Evaluation of Educational Achievement showed that U.S. students compared favorably with their peers in other nations. [220:20-21]

Christopher Jencks, professor of sociology at Harvard University, maintains that young people's test scores today are no better and no worse than their parents' test scores in the 1940s. Only if the trend is extrapolated into the distant future does it appear alarming, Jencks said. "The decline in test scores does not prove that students are stupider than students ten or fifteen years ago. They may have exactly the same level of aptitude but be directing it to other activities." Further, since schools do not teach aptitude, SAT scores are not useful in assessing whether high schools are doing their jobs. [141:13] Jencks believes that the only way to judge the significance of the decline in test scores is to examine the test item by item to ascertain whether each question is actually important. [23:1-2] Rather than trying to restore respect for authority figures such as teachers and parents via the back-to-basics movement, he suggests that schools should be trying "to restore respect for the value of reason, in all its complexity." [141:14]

In April 1979 testimony before the Senate Subcommittee on Education, Arts and Humanities, Roger Farr, president of the International Reading Association and Professor of Education at Indiana University, stated that fluctuations in reading achievement data could be "traced to such factors as changing promotion policies of schools, the move toward open enrollment policies, an increased emphasis on encouraging minority populations to seek more education, and a definite reduction in school dropouts." He commented, however, that these fluctuations made it seem as though a change in reading achievement was occurring, but what was really happening was a change in the composition of the test-taking population. He testified that the SAT was not a valid assessment of reaching achievement in the public schools because it measures a very high level of comprehension, not functional literacy. Since the SAT is administered to college-bound high school juniors and seniors, Farr said an extension of the interpretation of scores on the SAT to attack reading education in the United States was "a gross misuse of such tests as well as a patently invalid argument." [86:2, 5]

In October 1979, Shirley Hill, president of the National Council of Teachers of Mathematics, testified before the House Subcommittee on Elementary, Secondary and Vocational Education. According to Hill, many schools are now teaching objectives that can be easily taught and tested. She testified that many teachers have "settled down to a single-minded dedication to one goal--high scores on tests of minimal skills." Hill added that "short-term retention is the goal, not long-term retention and the ability to apply." [84:2]

Patricia Lund Casserly, a research scientist at ETS, reported that traditional attitudes and behaviors of guidance counselors, teachers, and parents toward females result in young women not being adequately prepared in high school to pursue college courses in mathematics, chemistry, and physics, and to enter related professions. Casserly examined curriculum and guidance policies at 13 high schools nationwide and approximately 200 girls participated in the study. She emphasized that math teachers should encourage young women to persist in mathematics courses. "We don't let students drop English because they have trouble spelling, or because their grammar is imperfect or because they don't enjoy the selected literature," Casserly said. "Let's not do it in mathematics either." [289:7]

According to Gordon Cawelti of the Association for Supervision and Curriculum Development, recent declines in SAT scores raise questions about the effectiveness of competency-based education. "The competency movement has been the political response to the back-to-basics movement," Cawelti said, "and schools have been preoccupied with doing needs assessments and testing at the expense of focusing on direct instruction." Research evidence, he added, indicated that no significant increase in achievement can be anticipated without an increase in direct instructional time. Cawelti called for curriculum validation studies to determine the extent to which the high school curriculum teaches the verbal and mathematical reasoning abilities measured by the SAT. [241:2]

According to W. Timothy Weaver from the School of Education at Boston University, there is evidence of declining standards and much lower SAT verbal and mathematical scores among entering freshmen in teacher education compared to entering freshmen in other fields. In 14 of the 19 postsecondary institutions Weaver surveyed, there was a significant decline in the quality of applicants who intended to major in education as college-bound seniors shifted career choices to those in greater demand. [290:568] At eight teachers' colleges the verbal test score of freshmen fell from 472 in 1970 to 417 in 1975, which exceeded the national decline in SAT scores. [290:570-571] In 1975-76 the SAT verbal score of college-bound seniors who planned to major in education was 34 points below the national mean, 397 compared to 431, and the SAT mathematical score was 43 points below the national mean, 429 compared to 472. Education majors were reported to be second from the last on the ACT mathematics test and tied for fourteenth place on the ACT English test in 1975-76, compared to students in 19 other career fields. [290:575] Weaver said the decline in SAT scores takes on added significance in education because "if members of this group constitute most of those who will apply to schools of education, then they also constitute most of those who will be accepted." [290:585]

In 1979 the National Institute on Drug Abuse conducted a nationwide survey of drug use among approximately 17,000 high school seniors in 130 public and private schools. The study, *Drugs and the Nation's High School Students. Five Year National Trends*, found that the class of 1979 had more experience with marijuana than any other previous cohort in American history. Other salient findings were the following: (1) the percentage of all seniors using marijuana at least once in their lifetime rose from 47 percent in 1975 to 60 percent in 1979. [143:23] During the same period daily or nearly daily usage of marijuana increased from six percent in 1975 to 11 percent in 1978; it remained unchanged at 11 percent in 1979; [170:1] (2) there has been a dramatic rise in the annual prevalence of cocaine usage with the percentage of users doubling from 5.6 percent in 1975 to 12 percent in 1979; (3) the annual prevalence of inhalants increased from 3.0 percent to 5.4 percent, but this is considered to be an underestimate since users often fail to report amyl and butyl nitrates under the inhalant category; (4) the annual prevalence of stimulants rose from 15.8 percent in 1976 to 18.3 percent in 1979. On the other hand, annual prevalence of sedatives, tranquilizers, and heroin declined among seniors. Sedatives declined from 18.2 percent in 1975 to 14.6 percent in 1979; tranquilizers, from 10.8 percent in 1977 to 9.6 percent in 1979; and heroin, from 1.0 percent in 1975 to 0.5 percent in 1979. [143:23, 28] According to Sidney Cohen, professor at the University of California at Los Angeles, a "juvenile stoned on pot or any drug during his waking hours" results in lost "learning time, growing-up time, problem-solving time." [169:4]

At a January 1980 meeting of the Senate Judiciary Subcommittee on Criminal Justice, Harold Voth, psychiatrist at the Menninger Foundation in Kansas, testified that marijuana smokers exhibited "loss of motivation, lowering of ambition" and desire "for excellence." H. Brian Berthiaume, program coordinator of the Phoenix School in Montgomery County, Maryland, testified that listlessness, poor attention spans, and lower grades of children who smoke marijuana have "serious ramifications for the classroom." Berthiaume noted that students had trouble with complex tasks and become, "frustrated with difficult math problems." Senator Charles Mathias (R-Md.), who chaired the Senate Subcommittee on Criminal Justice hearings stated that drug usage might contribute to curriculum problems. [78:3]

The President's Commission on Foreign Language and International Studies reported a "serious deterioration in this country's language and research capacity." The Commission's report, released in November 1979, showed that only 15 percent of American high school students now study a foreign language compared to 24 percent in 1965; only one in 20 students takes courses in French, German, or Russian beyond the second year; and only eight percent of American colleges now require a foreign language for admission, compared with 34 percent in 1966. Commission chairman James A. Perkins called upon schools, colleges, and universities to reinstate foreign language requirements. [219:5-7]

A professor of radiological physics at the University of Pittsburgh School of Medicine has claimed there is a connection between low SAT scores in the 1970s and nuclear bomb tests in Nevada in the 1950s. [265:184] Ernest Sternglass said that radioactive fallout from atomic bomb blasts hindered the normal development of thyroid and pituitary glands of fetuses. [242:A10] He cited a study which showed a three- to fourfold increase in thyroid conditions among young people in Utah 10 to 20 years after the fallout of the mid-fifties. [265:185] Sternglass also observed that the largest drop in SAT scores during the two-year testing period, 1973-74 and 1975-76, occurred in the western region of the United States, where fallout was heaviest and mixed with rain from Russian tests in Siberia and American tests in the Pacific Ocean and Nevada. Using regional data obtained from the College Board, Sternglass noted that SAT scores declined 19 points in the western region, compared with 12 points in New England, 14 points in the Middle Atlantic States, and 13 points in the South. Children in Utah experienced a decline of 26 points, in contrast to children in Ohio who experienced a two point decline. [265:184] Other areas of the nation which received significant amounts of fallout experienced significant dips in SAT scores, Sternglass asserted. [242:A10]

Although there is considerable controversy surrounding the interrelated factors which may have contributed to declining test scores, the College Entrance Examination Board has not published any additional studies on the reasons for declining test scores since its 1977 report, *On Further Examination*. [161:1]

THE TESTING CONTROVERSY: RALPH NADER VS. ETS

Until recently, almost all criticisms of tests and testing initiated within the psychometric profession, according to William W. Turnbull,* former president of ETS. As the costs of education have risen and the school population has declined, many consumers of education have not seen the increases in the quality of education that they expected. Consequently, the public has demanded accountability and improved pupil performance, which has led to an increased questioning of the educational process in an effort to understand why student achievement has not matched anticipation. [283:105] This questioning, however, has turned into an "open assault" on the college admissions process and institutional autonomy, in the opinion of George H. Hanford, president of the College Board. [273:4] Consumer groups, public interest groups, and some teachers' organizations are, in effect, "testing the tests." [47:2; 123:1] Turnbull has indicated that the national debate about standardized testing has shifted substantially from one of inquiry to a declared "*war against* testing." (emphasis in the original) He has characterized these most recent attacks as "among the most sustained and hostile assaults on an applied social science ever seen in this country." Because the media and political arenas have become battle sites, Turnbull believes that many untruths and misleading statements have gained popular acceptance without a complete examination of the facts and their possible consequences for all concerned. [68:2]

According to Turnbull, the attacks upon standardized testing appear to focus on the possible misuse of test results and the role of ETS as one of the nation's most prestigious testing organizations. [68:2] But there are a host of subissues. Foremost among these concerns are whether the SAT measures aptitude and predicts first year college performance; whether the tests are culturally biased and exclude a disproportionate number of minority applicants from college; whether the test scores relate directly to family income and perpetuate a "class system in the guise of merit;" and whether the testing organizations are accountable for products which affect the lives, careers, and aspirations of millions of Americans each year.

The continuing controversy over standardized testing peaked in early 1980 with the release of a Ralph Nader sponsored report on ETS titled *The Reign of ETS: The Corporation That Makes Up Minds*. [188] One educational observer speculated that the Nader study could eventually affect not only the future of the College Board's SAT tests, but the future of all standardized testing, including tests given at the elementary school level. [173:A19]

---

*ETS president William W. Turnbull resigned in October 1980 and plans to study the testing process as a resident scholar at the ETS facility in Princeton. He has also called for wide debate on testing for college admissions.

Origins of the Nader-ETS Feud

The feud between Ralph Nader and ETS extends back to 1974 when a Nader associate, Allan Nairn, then a high school graduate, began an investigative study of the claims ETS made about its standardized tests. [104:5] Nader formulated the concept for the report as he toured colleges and universities across the country. Students approached him to complain that, in spite of their grades and extracurricular activities, they had been wrongly judged on the basis of a three-hour examination. Nader reported that many students felt that their educational and career opportunities had been irreparably damaged or destroyed by their low test scores. Because these students did not "test out," they believed they lacked the aptitude to go to graduate and law school, they exhibited an attitude of resignation, and they viewed the results of these multiple-choice tests as "revealed truth about themselves." [188:ix]

In response to the students' complaints, Allan Nairn and others, under Nader's sponsorship, formally approached ETS officers for interviews and published materials. [188:xi] Nader claimed that ETS executives demanded financial compensation for time given in interviews, retention of a court reporter to take transcripts, a pledge that the information obtained would not be used in litigation, and the power of reviewing the study prior to its publication. [188:xii] Nader said that previous investigative studies he had conducted of General Motors, DuPont, and Citibank, were "child's play" compared to ETS and its "seige-mentality." [171:23; 188:xiv]

In rebuttal to Nader's statement, ETS's Turnbull said that when it became apparent that Nairn wanted to conduct an intensive, time-consuming study, ETS agreed to participate, provided that a systematic set of procedures would be followed. Accordingly, on two separate occasions, Allan Nairn and Thomas Sutton contacted ETS and were given well over 100 publications and access to the ETS library. Six weeks after their second investigation began, however, the researchers unexpectedly disappeared without completing all of the interviews planned. Turnbull reported that there was no explanation from either Nader or Nairn for this action. [284:1-2]

In January 1980, after five years of work, Nairn, now a 24-year-old Columbia University economics student, and his associates released their 554-page study, *The Reign of ETS: The Corporation That Makes Up Minds*. In the foreword to the report, Nader criticizes ETS for its extraordinary power of deciding who advances through what school, a power that would have created a furor years ago if a governmental agency had possessed such authority. [188:xvi] According to Nader, the report shows that ETS's claims for its aptitude tests are insupportable. [188:xvii] Further, Nader says that a "one-time three hour gamble which can determine a life's pathway is simply not compatible with what is known about human personalities, their capacity for growth, their diversity and versatility." [188:xv]

Calling the Nader and Nairn report on ETS and standardized testing an "anti-climax after a five-year build up," Turnbull stated that much of the material was dated and its conclusions were erroneous. He emphasized that ETS had tried to correct the record repeatedly but that the corrections had been ignored and inaccurate statements repeated. Turnbull accused Nader and Nairn of confusing monopoly with success, blaming the tests for showing that minority students are inadequately prepared compared to majority students, and ignoring the fact that standardized tests reduce the possibility of unfairness in the admissions process. By attempting to shake the public's confidence in ETS and standardized testing, Turnbull claimed that Nader was trying to eliminate standardized testing altogether and substitute college admissions standards with his own, as yet, undetermined subjective values. Also, Turnbull accused Nader of being out of touch with educational needs and public sentiment concerning testing. [285:1-3]

According to Turnbull, the major criticisms of ETS presented in Nairn's report appeared in the May 1975 issue of *Ladies' Home Journal*, which indicated that a verdict had been reached long before ETS' "trial" started. [285:1] The report, said Turnbull, willfully ignored many of the new initiatives ETS had recently undertaken. He suggested that a reasonable perspective toward standardized testing should be taken to improve both the tests and their use. [68:2] "If testing is weakened," Turnbull cautioned, "education and society will be the losers." [285:3]

Major Issues in the Debate

Highlights of 17 major issues debated between Ralph Nader and ETS, and between testing critics and testing advocates follow.

The nation's gatekeeper.-- Allan Nairn has charged that ETS is "the gatekeeper to educational and career opportunities" in the United States. [188:292] In its role as the nation's gatekeeper, ETS, in effect, determines who can attend college, graduate school, and professional schools, as well as who can be certified as a teacher, auto mechanic, plumber, and beautician. [184:2; 171:22] By claiming it has evolved the "science of mental measurement," ETS protects and enhances its role as the gatekeeper; however, its tests actually measure little more than how a candidate responded to a few multiple-choice questions, Nader asserted. [184:2] Similarly, Benjamin L. Hooks, president of the National Association for the Advancement of Colored People, charged that ETS's testing instruments which certify, classify, and stratify, "have effectively and disproportionately screened out blacks and other minority" candidates, thereby limiting their access to employment, colleges, and professional schools. [131:1]

ETS, on the other hand, challenged Nairn's allegation that it is the "arbiter of admissions" or "gatekeeper" to higher education in America. ETS articulated that colleges, graduate schools, and professional schools make their own decisions regarding which candidates to admit based on each institution's own criteria. These criteria are determined by administrators and faculty at these institutions and by state boards of education. [77:11] Because the selection process varies considerably from institution to institution, there is no single "gate" in the admissions process. According to ETS, its standardized tests serve as a means of testing in, rather than as a means of testing out, of opening gates to opportunity, not of closing them. During the past two decades, ETS reported that the expanded use of tests has accompanied an increase in the proportion of minority and women candidates in institutions of higher education. [65:7] George H. Hanford, president of the College Board, stressed that between 1970 and 1977 enrollment rates for blacks and Hispanic candidates increased by five percent and six percent, respectively. [121:5] Turnbull added that it would be hard to infer that admissions tests have barred college doors to minorities and the disadvantaged. [68:4]

Monopoly power.-- According to Nairn's report, ETS has a complete monopoly in eight of its top ten testing markets. [188:260] Nader characterized this situation as unprecedented in corporate history. [184:1] Terry Herndon, Executive Director of the NEA, reported that ETS controls a key 14 percent of the total standardized testing market, and that its yearly gross income is nearly four times as large as its nearest competitor. [126:2] Roger Lennon, senior vice-president of Harcourt, Brace, Jovanovich, Inc., stated that ETS enjoys a privileged advantage over its rivals in private industry since "academic types tend to feel more at home dealing with something that isn't perceived as a commercial enterprise." [300:18] Over the past 30 years, it has been estimated that ETS's standardized tests have impacted upon the lives of 90 million people. [126:2]

ETS asserted that Nader and Nairn confused monopoly with success, and further, that schools and colleges are free to choose among a variety of nonprofit agencies, government agencies, and private companies to carry out their testing programs. While ETS is a major element in the testing field, the company states that it is proud that so many institutions use its services because the standardized testing market is so highly competitive. Other commercial enterprises like CTB/McGraw-Hill, Psychological Corporation (a division of Harcourt, Brace, Jovanovich, Inc.), Science Research Associates (a division of IBM), Houghton Mifflin, and the Measurement Research Center (a division of Westinghouse Corporation), have financial resources similar to those of ETS. [65:3; 70:2]

Forced consumption.-- Nairn stated that test takers are "captives" who have little, if any, choice: as involuntary consumers they have no say in determining whether they will pay and take admissions tests. [188:262] Currently, the SAT is required by 50 percent of the nation's colleges, the Graduate Record Examination (GRE) by 75 percent of the graduate schools, the Graduate Management Admissions Test (GMAT) by 80 percent of business schools, and the Law School Admission Test (LSAT) by 100 percent of law schools. [238:4A; 187:58G] In short, most students must take tests developed by ETS, or face abandonment of their educational plans. Some see this as a double jeopardy situation for, as one parent commented, the candidate is obliged "to generate evidence which may, and very often is, eventually used against him." [188:261-262]

In contrast, Educational Testing Service emphasized that candidates who register with ETS to take an admissions test seem to associate the testing organization with the test requirement. But the requirement actually comes from the institution to which the candidates are applying, or from licensing and certification agencies that establish their own criteria. Moreover, many colleges and agencies exempt candidates from tests or use other methods of evaluation. ETS said its so-called power over candidates is a myth because it merely provides the tests, and has no influence on whether or not the candidates will be tested. [65:7]

Accountability.-- Although ETS maintains that it is accountable to its client groups, these groups are not free to choose among different competitors for services, Nairn stated. With the exception of the College Board, ETS's other client groups were *created* by ETS to sponsor programs it already owned. Specifically, Nairn said the LSAT and GMAT client boards were created by ETS in 1953-54. Other client boards have been created since then, such as the Graduate Record Examinations Board in 1966. [188:303] "Although the client groups enjoy varying degrees of autonomy," Nairn wrote, "whatever authority they hold is granted under terms set by ETS as the owner of the test questions." [188:304] Nairn also reported that as of 1974, the New York State Board of Regents, which granted ETS its charter in 1947, had not communicated with, made any inspection of, nor required any submissions from ETS in its 27-year history. [188:282]

ETS reiterated that its clients are independent and that ETS is fully accountable to them.

Each client organization contracts with ETS for admissions testing and related services. Also, the decision to use ETS's services resides with the educational institutions and their representatives. ETS noted, for example, that from 1947 to 1960 it developed and administered the Medical College Admissions Test for the American Association of Medical Colleges (AAMC), but since 1961 AAMC has contracted this work to other testing organizations. As a nonprofit educational organization chartered under New York's education laws, ETS said that it is accountable to the New York State Board of Regents. [65:3] Furthermore, like all nonprofit organizations, ETS is accountable to the Internal Revenue Service. On a professional level, ETS is accountable to its Board of Trustees and its staff. [70:1]

Nonprofit status.-- Nairn reported that ETS obtained its nonprofit charter from the New York State Board of Regents, and that the company was granted an exemption from federal income taxes under section 501 (c)(3) of the Internal Revenue Code. [188:299-300] Nairn suggested, however, that ETS bears little economic similarity to the majority of nonprofit civic groups, churches, hospitals, and schools because ETS is an enterprise "selling consumer products on a national scale." Since ETS has no shareholders, it is immune from the risks associated with public ownership, such as responding to stockholder suggestions or complaints. "Its freedom from antitrust jurisdiction facilitates the establishment of client relationships," Nairn commented, "which help insulate ETS from the marketplace of consumer choice." [188:298-299] Yet, in spite of ETS's nonprofit status, Nader said ETS declares approximately $1 million in "nonprofits" each year. These funds are used for corporate expansion and to maintain the ETS campus: a 400-acre headquarters, a $250,000 home for the president, and a $3 million hotel/conference center; all constructed with candidate test fees. [184:1] Although ETS has only a small revenue of stocks and cash, Nairn stressed that ETS has an overwhelming political and economic interest in preserving its testing system. [188:395]

In response, ETS stated that its nonprofit, tax-exempt status is similar to thousands of religious, consumer, and charitable organizations which state and federal governments have deemed to be in the public interest. [70:2] According to ETS, the bulk of its income in 1978-79, some $85 million, came from services it provided for test sponsors such as the College Board, graduate and professional school testing programs, and other testing activities. Almost $8 million came from outside sources such as local, state, and federal agencies, as well as from studies and activities funded by foundations. An additional $1.5 million came from fees at the ETS conference center and returns on invested reserves. [65:4] It was noted that ETS publishes an annual report, containing a complete financial statement, which is sent to the educational community, the media, members of state legislatures, and the general public. [70:1]

In 1978-79, ETS reported that its excess income over expenses amounted to $1.1 million or 1.2 percent of expenses. This money was used for acquisitions, modifications of operational space, capital equipment, and working capital. ETS said that the company pays taxes on "unrelated income" as determined by IRS and voluntarily pays property taxes on its lands and buildings in Lawrence Township, New Jersey. Although ETS stated that it maintains only a small invested reserve, less than one month of its annual payroll, the organization says that its nonprofit status should not be construed to mean it pursues a strategy of no growth. [65:3; 70:2]

Center for Occupational and Professional Assessment.— According to Nairn, as the number of college applicants began to plateau in the 1960s, ETS directed substantial amounts of its promotional effort to the occupational testing field, following students out of the classroom and into the professions. [188:325; 300:1] Since 1969, occupational testing revenues have grown fivefold, from $2.1 million to $11.2 million. [188:325] A partial list of the occupations for which the Center for Occupational and Professional Assessment (COPA) had developed and administered written performance measures and/or program work includes accountants, bankers, real estate brokers, automobile mechanics, beauticians, plumbers, urban planners, electrical contractors, police officers, Peace Corps volunteers, Foreign Service officers, Central Intelligence Agency workers, and National Security Agency members. [300:1; 283:104; 65:2]

Turnbull stated that the formation of COPA was in response to consumer demands for competence and certification of various professions. Working closely with representatives of the sponsoring organizations and content specialists, COPA has developed assessment measures for many occupations and professions, particularly in health care. COPA has certification tests for gynecologists, obstetricians, opticians, respiratory therapists, nurses, opthalmologists, and pharmacists. Turnbull also reported a trend toward self-assessment programs which provide professionals with the opportunity to assess themselves in relation to their peers. Professional groups such as the American College of Dentists, the American Psychiatric Association, and the American Academy of Pediatrics have participated in such programs. [283:104-105] Now completing its sixth year, COPA's new projects include licensing and certification examinations for construction code inspectors in the electrical, building, plumbing, mechanical, and fire protection fields; a continuing education program for physicians; and new certification procedures for social workers, which combine test scores, references, and experience into a composite score. [68:11]

ETS data banks.— ETS has the largest data bank of personal, educational, and psychological information in the world, Nairn declared. Its files contain information on more than 32 million persons from 100 nations. [188:28] Testing critic Steven Levy reported that ETS has information on more than 15 million Americans, not only their test scores, but also the most detailed forms about personal finance that many people will have to complete in their lifetimes. Levy stated that it is on "trust" that ETS protects this information for the public. [166]

ETS acknowledged that it maintains computer records of the test scores and other data about millions of candidates, but it stressed that this information is provided voluntarily by the candidates. Moreover, this information is completely confidential. It cannot be released to an institution or agency without the specific permission from a candidate. The data, however, are available to qualified researchers under rigorous professional guidelines which assure confidentiality. ETS noted that its intensive concern for confidentiality has lead, ironically, to allegations that ETS operates under a "veil of secrecy." ETS stated that it recognizes the rights of candidates and institutions to privacy with regard to information supplied by and about them, as well as ETS's responsibility to guard information in its files from any unauthorized disclosure. [65:2, 4]

Test development.— In 1978 it was reported that the SAT was written and updated by a staff of 58 test developers, which included ETS's own staff and a sizeable free-lance pool. [300:18; 85:41] But the question writing process, according to Nairn, involves no formulas or statistics. [188:144] He reported that it takes anywhere from a half-hour to an hour to write a typical SAT verbal question, which is reviewed at several stages by two to five ETS staff members or consultants. [188:147] Questions which survive the review process are then pretested to establish whether they are "easy" (more than 70 percent get it right) or "hard" (fewer than 30 percent get it right). If more than 90 percent get a question right or wrong, it is eliminated. [300:18] Although ETS executives maintain that 150 steps are involved in the test development process, Nairn said that the majority of these steps are office procedures: ("Step No. 65. Prepare Official Key for Reproduction, 66. Proof Key and Release for Printing, 67. Key Sent to Printing, 68. Print Key, 69. Distribute Key to Appropriate Divisions. . .") [188:147] Nairn said that over the past 30 years, ETS has accumulated an inventory of approximately 300,000 reusable multiple-choice test items, some of which have appeared in as many as 12 different tests. [188:145-146]

ETS's Turnbull stated, on the other hand, that test development is conducted by ETS staff specialists who work in conjunction with representatives from the organizations and associations for whom a test is being developed. ETS staff specialists are aided by a committee of content area experts. All examinations, in turn, are reviewed, pretested, and analyzed by ETS staff specialists, an organization's program policy board, and field experts. Additionally, ETS asks for advice from groups representing various disciplines and professions and requires minority representation in the review process to eliminate any possible cultural or racial bias. [283: 106] Before a single question appears on a test, it will undergo nearly 30 inspections, ETS asserted. [85:41] Because of the screening process and pretesting, it can cost $100,000 or more to produce a new Graduate Record Examination or a new Law School Admissions Test, and it can take as long as 24 months to prepare a single question included on a final test. [85: 41; 188:147-148]

Standard error of measurement.--The standard error of measurement on a SAT score extends from 30 points below a candidate's actual test score to 30 points above, Nairn explained. This means that if a candidate takes the SAT an infinite number of times, the "true score" or the average of all the scores would fall within this 60 point range two-thirds of the time, while a third of the time it would fall outside of this range. [188:157] Steven Levy said that, in theory, this means a student with a score of 540 should be considered identical to a student scoring 600. [166] In 1971 researchers Wing and Wallach analyzed data from 224 colleges of varying degrees of selectivity. They found that at many of the schools a score difference of *less* than 60 points could alter chances for admission by 100 percent. Moreover, a score increase for an applicant from 425 to 475 could *raise* the chances for admission by 70 percent. A decrease of 50 points, however, would *lower* the chances of admission by an equal amount. From this research, Nairn suggested that the test points which can alter a candidate's educational future are "derived from statistical formulas that transform the answers to a few multiple-choice questions into an elaborate three-digit report of 'aptitude.'" [188:156-158]

Turnbull noted that while test scores as well as other sources of information about human measurement and performance contain error, the error of measurement on standardized tests is well known and announced by publishers. According to ETS, this consistency provides a contrast to essay examinations, which many studies have shown to be unreliable, since grades assigned by readers vary considerably. Furthermore, unlike other forms of measurement such as teachers' grades, letters of recommendation, and personal interviews, test scores are designed to provide a common basis for the evaluation of a candidate's abilities to perform well in future academic work. [77:8; 69:4]

Aptitude.— The earliest founders of mental measurement, including Carl C. Brigham, father of the modern SAT, emphasized the relationship between test scores and social class, Nairn reported. These early psychometricians believed that groups which scored low on tests were intellectually inferior. [186:652] Writing in *The Atlantic*, James Fallows reported that early mental testers administered IQ tests to newly arrived immigrants at Ellis Island in New York City, which resulted in the exclusion of certain racial and ethnic groups. For example, in 1912 Henry Goddard "scientifically proved" that 90 percent of Hungarians, 87 percent of Russians, 83 percent of Jews, and 79 percent of Italians were "feebleminded." [85:39] Today, Nairn claims that ETS views "aptitude" in the same way that the early mental testers viewed "intelligence." Nairn accused the psychometric profession of assuming a role beyond objectively reporting how people perform on multiple-choice tests and of presenting the tests as scientific findings which purport to measure the basic quality of an individual's mind. [188:383] Similarly, two Harvard Medical School researchers, Warner V. Slack and Douglas Porter, stated that "aptitude, like intelligence, implies native mental potential." Therefore, the application of the word "aptitude" to any measurement based on test scores should be viewed with caution. [258:169]

Slack and Porter emphasized that ETS promotes the SAT as a measure of aptitude and that the public assumes high test scores indicate students have good minds and, therefore, benefit from favorable judgments in comparison to those who have low test scores. Nonetheless, Slack and Porter stated that the SAT has not been subjected to serious scrutiny because college admissions committees lack the resources and time. In fact, admissions committees are relying upon test scores to help expedite the time-consuming job of selecting and rejecting applicants. According to Slack and Porter, "if aptitude is defined simply as the capacity to do college work, the SAT has less relevance than either high school grades or other achievement tests." [258:172] They contend that there is no evidence that the SAT measures anything except learned skills, for which training is effective. They pointed out that ETS's achievement tests in mathematics, science, and language actually have higher correlations with college performance than the SAT. Moreover, the English composition test is the only subject for which the SAT predicts as well. Slack and Porter referred to the SAT as another achievement test: "If stripped of the aura of 'aptitude' and considered together with other achievement tests and high school grades, the SAT is a third-rate predictor of college performance." [258:171] They called upon colleges and universities to follow the example of Bowdoin College and eliminate the SAT requirement. They also reminded students who are disappointed with their test scores that the scores do not reflect "the quality of their minds." [258:172] Ralph Nader added that it is presumptuous to define what aptitude is, let alone to allocate educational opportunities based on it. Such notions should be directly challenged, he said. [188:xvi]

One of the most unfair sections of Nairn's report, according to George Hanford, president of the College Board, tries to link current testing procedures with earlier mental testing experts whose views on racial superiority and eugenic sterilization have long been discredited. [121:2] Hanford also accused Slack and Porter of ignoring the College Board's description of the

SAT, that is, the SAT is "a test of developed abilities--a measure of wide-ranging learning experience--and that it is not a content-specific achievement test in the conventional sense of that term." [122:1]

With regard to IQ tests, Turnbull stated that ETS neither defines intelligence, develops intelligence tests, nor reports IQ scores. [284:2] ETS's aptitude tests, moreover, were never intended to assess innate intelligence or unchanging abilities. Rather, they measure learned or acquired skills which develop slowly over time through school and nonschool experiences. These tests are described as aptitude tests since they are not tied to a specific course of study, school curriculum, or program. [77:7] ETS asserted that it is constantly seeking ways to improve the evaluation of abilities related to success in academic pursuits, develop tests of these abilities, and validate them using the most rigorous psychometric techniques currently available. [65:6] While these measured skills are applicable to success in a variety of future academic endeavors, Turnbull commented: "To think that, at 18 years of age, people whose experiences have been vastly different can show their inborn potential through a test of verbal and mathematical reasoning is naive, regardless of their cultural advantages or disadvantages." [69:4]

Self-worth.-- An individual's test scores, Nairn stated, often become the basis for important changes in self-perception and self-concept. [186:653] He cited several studies which document the effects of test scores on candidates' aspirations. In a six-year study of 80,083 students sponsored by the College Board, Dale Tillery of the University of California wrote that "while in high school, students learn to judge themselves by these same measures [scores and high school grades] and, when they do not measure up, they either abandon certain educational and career goals or try circuitous routes to achieve them." [188:216] Another study, conducted by Dennis Dugan, focused on the aspirations of students attending 47 high schools in the Boston area in 1969. Dugan's study found that the student's SAT score "has its greatest impact" on his or her decision to "enter the labor market. . .instead of pursuing more education." [188:216-217] Nairn pointed out that there have been no studies to estimate how many lawyers, doctors, and other professionals from the black, Latin, and other ethnic communities have been lost by relying on ETS aptitude tests. [187:62GE]

Similarly, Slack and Porter accused ETS of willful omissions in its technical publications which enhance the appearance of their product at the expense of the candidate's self-esteem. [258:172] Slack and Porter wrote that "if a student with a low SAT score believes that this score accurately reflects his aptitude, and that no study or preparation could have altered his performance, his mental capacity is disparaged, and his self-esteem is weakened." [258:155] The ultimate tragedy, according to Nader, occurs when students accept their test results as a measure of their self-worth. Nader went on to say that broader and more diverse approaches for assessing individual performance are essential in order to break this "vicious circle" and end the "reign of ETS." [188:xiv]

In rebuttal, ETS called Nairn's report a distortion of the record. Guidelines published in College Board bulletins and manuals for students, high school counselors, and admissions officers include statements such as the following:

- "Tests can be a useful measure of knowledge or academic ability, but no test predicts with any certainty 'success in life' or is in any way a measure of an individual's total worth."
- "Test scores, like all types of measurements, physical as well as psychological,

are not perfectly precise and should not be treated as though they were."

- "Although admissions test scores are good predictors of performance in college, they are not infallible predictors." [77:10-11]

ETS further reported that there is no real evidence to support the contention that taking a standardized test has a damaging psychological effect on a person. [65:7] Robert J. Kingston, former president of the College Board, stated that standardized "tests are only one piece of evidence used in one professional context." Kingston added that standardized tests reveal nothing about character, promise beyond college, chance of succeeding in life or "our ability to take our place as worthwhile human beings." [47:6]

Creativity.-- Nader and Nairn said that although standardized tests such as the SAT can determine a life's pathway, these tests cannot measure the variety of skills which are essential to success in any kind of endeavor. [188:xv; 391] Nader claimed that ETS's tests do not measure important human qualities which have advanced civilization, such as creativity, wisdom, judgment, determination, stamina, idealism, and experience. [104:5] Standardized tests simply measure the specialized skill of multiple-choice test-taking, he asserted. [184:2]

ETS's president acknowledged that Nader was correct in stating that the psychometric profession has no large scale tests for assessing human qualities such as creativity, idealism, stamina, and caring for mankind. Turnbull stated that those qualities are very important in society and that they should not be ignored in favor of dependence on tests alone. He pointed out that ETS has been actively seeking to extend the range of available measures useful in assessing a candidate's abilities. [284:1] Hanford observed that the SAT does not measure height, weight, skin color, geographic location, or economic status because the test is not designed to do those things. Hanford affirmed that the SAT remains the best "national yardstick of student ability...regardless of school, class, curriculum, family background, or other variables." [123:4]

Meritocracy.-- A British sociologist, Michael Young, coined the term "meritocracy" nearly 20 years ago to describe a system of rewards based on ability and merit "rather than accident of birth." James Fallows observed that Young's term was satirical in intent, but Americans, in their search for an equitable system of classification, appropriated it without its original irony. [85:43] Nairn said that many educators view the SAT as a standard of merit and, while recognizing that the SAT is less than perfect, they consider it an imperfect standard which is better than no standard at all. [188:392] Nairn said that this kind of argument fails to consider the availability of other measures. He said that ETS's own research has indicated that judging candidates by their previous *accomplishments* would almost *eliminate* class and ethnic discrimination factors which are prevalent in the ranking of candidates by ETS scores. [188:392] (emphasis in the original) Nairn stated that the test system itself 'routinely damages the poor, the working class and the middle class relative to the more affluent and privileged." [188:391] It also allows an administratively convenient way of avoiding the responsibility of judging candidates. [188:393] According to Fallows, standardized tests have merely exchanged one kind of privilege for another. [85:47]

According to the College Board, however, from the beginnings of the modern SAT, Carl C. Brigham called for a reasonable perspective on tests and their use. Brigham wrote that "to place too great emphasis on test scores is as dangerous as the failure properly to evaluate any score or rank in conjunction with other measures and estimates which it supplements." [47:3] Throughout the history of the SAT, and especially since World War II, the CEEB has viewed the increasing use of standardized tests as a democratizing influence in higher education. [47:3] By virtue of their accuracy, objectivity, and comparability, standardized tests have served to identify individuals who regardless of class, sex, or ethnicity, are most likely to perform successfully in college. [69:2] According to George Hanford, Nader and Nairn are trying to reorder the nation's college admissions process and current system for educational opportunity and substitute subjective and undefined values of their own for objective measures and a meritocratic approach. [121:2, 4] Hanford suggested that Nader's attacks in the guise of criticizing methodology and statistical accuracy are essentially attacks upon the concept of standards and quality in education and society. [121:2] The standards for college admission should be left to the colleges and universities, said ETS president Turnbull. [285:2]

Minorities.-- Nairn reported that an unbiased test according to ETS's definition is one that "predicts the first year grades of minorities about as accurately as they predict the first year grades of whites." But this definition, Nairn charged, fails to recognize that the validity of grade prediction is low for *both* groups. [188:112] (emphasis in the original) Contrary to the belief that admissions standards are lower for minority candidates, Nairn said that they must earn *higher* grades than whites in order to have an equal opportunity for admission to college. [188:111] (emphasis in the original)

Nairn cited a 1971 study of black and white students at integrated colleges conducted by ETS researchers Junius A. Davis and George Temp. They reported that "while the SAT score means for (admitted) blacks were lower than those for their white counterparts, the mean high school ranks were higher." Nairn also cited another ETS study of the grades and test scores for all applicants to ABA accredited law schools in 1976. This study found that 68 percent of the whites and ethnically unidentified applicants with college grade point averages of 2.75 or above were admitted to law school, compared to 58 percent of the black applicants with identical averages. Among applicants with averages above 2.50, 64 percent of the whites and unidentified students were admitted compared to 51 percent of the blacks. Since the black applicants' grades ranked as high as the whites, Nairn said that the blacks were placed at a disadvantage by the results of their ETS aptitude test, not by their past performance. [188:111]

Nairn asserted that ETS maintains that the low scores of minorities reflect deficiencies in the preparation of minority candidates, and not deficiencies in the tests. [188:111] He reported that the low average scores of minorities were primarily a reflection of the tests' tendency to rank people by family income. That is, students from high-income families tend to receive higher test scores. [188:113] Nairn emphasized that it was one thing to tell people that they have been victims of less than adequate education, but it was something else "to use those scores to *prevent individuals* from pursuing opportunities." [188:117] (emphasis in the original) While ETS purports to illuminate educational inequality, Nairn said, in effect, ETS's aptitude tests actually continue it. He noted that as long ago as 1969 the Association of Black Psychologists called for a moratorium on standardized testing because these tests were the beginning of a downward spiral that deprived black applicants opportunities for advancement. [118:117-118] More recently, in January 1980, Benjamin L. Hooks, president of the National Association for the Advancement of Colored People, issued a statement calling on Congress to enact truth-in-testing legislation to regulate the testing industry. Hooks said that regulation was necessary because the testing industry, for the most part, determines who goes to college and professional school, who can teach, and who can be employed. "Our nation can ill afford to allow any private agency to limit the opportunity of a child to an education," Hooks stated, "which allows him/her to reach his/her potential nor to, through nonregulation, foster a system or process which penalizes the working class." [131:1]

While Turnbull acknowledged that black and Hispanic candidates consistently achieve lower average test scores than whites, he emphasized that it would be erroneous to jump from that reality to the conclusion that the tests were biased. In essence, this would be "blaming the messenger for the message." [68:3] Furthermore, ETS said countless validity studies have shown that the SAT, LSAT, GMAT, and GRE aptitude tests predict as well for minority and majority candidates, and admissions officials find test scores useful in making comparisons among minority students. [75:2]

Substantial data also show that disadvantaged and minority children do not receive equal opportunity in American education, Turnbull added. He noted that this deficit increases with the passing of time. But "the tests do not create the inequality; they reveal it." [68:3] Eliminating the tests would probably not harm candidates with high grades from well-known schools, but it would limit the chances of many minority and rural students whose abilities show up on the tests, but who have mediocre grades. [285:2] Likewise, ETS reported, there is ample evidence that these standardized tests help to identify talented candidates in inner city schools, new schools, nontraditional schools, and rural schools. College admissions officers often cite the usefulness of the tests in locating students who have not had the advantages and visibility or more affluent middle-class students. Standardized tests also help to eliminate "bias and inequalities inherent in grading systems, interviews, and personal recommendations," according to the testing company. [75:2]

Since the end of World War II, minority groups and people of low socioeconomic status have gained increased admission to college. Moreover, a high percentage of all candidates for college admission are admitted to the college of their choice. Turnbull said that, extrapolating from this trend, it would be difficult to conclude that standardized tests have barred college doors to minority students and the disadvantaged. [68:3-4] According to Hanford, the Nairn report implies that minority candidates are routinely being denied admission to college. On the contrary, said Hanford, between 1970 and 1977 the enrollment rates for whites remained fairly constant, while the "enrollment rates for black and Hispanic students rose by five percent and six percent, respectively. [121:5] ETS added that the evidence presented to the U.S. Supreme Court in the Alan Bakke case clearly indicated that test scores are used to identify potential talent and that minority candidates with above average, average, and below average test scores are being accepted into prestigious law and medical schools. [75:2]

ETS further criticized Nairn's report for not devoting attention to the role of the College Board in advancing the concept of awarding financial aid based on need, ignoring ETS's financial aid need analysis program at the graduate and professional school level, and overlooking many talent search, guidance, scholarship, and demonstration projects designed to ameliorate educational opportunities for disadvantaged and minority candidates. [76:10] ETS affirmed that it is committed to addressing the "causes of differences in educational achievement" for both the disadvantaged and the affluent. ETS reiterated that Nairn had ignored important research activities which involve ETS such as the following: (1) a five-year study of compensatory reading programs in grades 2, 4, and 6; (2) a longitudinal study of disadvantaged students and their initial school experiences; (3) studies of exemplary school desegregation practices; (4) evaluation studies on the effects of educational programs, such as *Sesame Street*, and the programs' influence on the skills and achievement of disadvantaged children; and (5) a number of evaluations of compensatory education projects conducted at the local level. [76:11]

Test scores and income.-- Standardized tests, Nairn stated, have not replaced ranking based on economic class with ranking based on merit. The so-called "democracy of multiple-choice tests," he contends, remains largely a ranking of people by family income. [188:198-199] Not only do the ETS test scores discriminate between the rich and poor, Nairn said, but they also discriminate between the rich and a *majority* of Americans, the working class and the middle class. [188:200] (emphasis in the original)

Nairn observed that "the more money a person's family makes, the higher that person tends to score." [188:200] The ranking by class is prevalent not only when large groups are averaged together but also among applicants to individual colleges. [188:202] This pattern appears to be consistent over geographic regions and income levels, Nairn added. A 1973-74 College Board report cited by Nairn showed, for example, that students with the highest SAT average (750-800) had an average family income of $24,124 compared to students with the lowest SAT average (200-249), $8,639. (See Table 7.) More recent data on the test scores and family incomes for candidates in 1978-79 revealed a similar pattern. For the applicants averaging below 350 had a mean family income of $18,400, while those averaging 650 or more had a mean income of $33,400. [188:201, 203]

Citing figures compiled by Humphrey Doermann for a College Board colloquium, Nairn reported that a student from a family earning less than $4,600 in 1969-70 income had a 10 percent chance of scoring above 450; the chances increased to 21 percent in the $7,500 to $10,699 range; and they quadrupled to 40 percent in the $16,200 range and above. [188:203] Nairn concluded that if ETS scores measure a person's "merit," then merit in the United States is allocated according to family income. [188:204]

ETS asserted that Nairn's report distorted the facts concerning test scores and family income. ETS said Nairn's allegations that the company had suppressed information concerning the relationship of test scores to students' family incomes, that SAT scores and family income rank students nearly the same way, and that the tests are used to preserve a social status quo, were all untrue. [76:6]

According to ETS, the data which Nairn used came from a series of reports, published by the College Board since 1971-72 and distributed to over 15,000 institutions and individuals. Although these data show that students from families with higher incomes tend to receive higher test scores, students from each income level obtained the full range of SAT scores. Also, nearly one-third of the candidates with family incomes below $6,000 ranked in the top half of the entire group in terms of SAT scores. [76:6] Thus, claims ETS the use of admission tests has contributed substantially to increased access of poor and working class students to higher education. [76:6] "It would be tragic if exaggeration of the relation between *average* income and *average* test scores were to lead to unfounded pessimism among able low-income students or their parents about their chances of doing well on tests," wrote ETS's Turnbull. (emphasis in the original) "And despite claims to the contrary, it is still a fact that test scores correlate more highly with grades than with family income." [68:4] Table 9 shows the average SAT scores and reported family income for college-bound seniors of 1973-74.

Prediction of first year grades.-- "The ability to predict grades," Nairn stated, "is the empirical basis of ETS's claim to measure aptitude." [188:58] ETS has maintained that countless studies have shown that the SAT is an effective predictor of college performance. Yet, Harvard Medical School researchers Slack and Porter contend that their research indicates otherwise. They reported that a summary of validity studies between 1964 and 1974 (Ford and Campos, 1977) indicated that the SAT validities for males and females combined averaged about .40 for the SAT-V, .35 for the SAT-M, .50 for the high school record, and .58 for all three predictors combined, i.e., the SAT provided an increase of .08 over the high school record alone. [258:165] Slack and Porter said that their own arithmetic, based on the data provided by Ford and Campos, showed that the SAT validity coefficients were smaller. [258:166] Table 10 shows the median validity coefficients of combined-sex samples for the SAT-V, SAT-M, high school record, and combined predictors, 1964-74 as reported by Slack and Porter.

According to Slack and Porter's analysis of the 1964-74 data, the validity coefficients are .37 for the SAT-V, .32 for the SAT-M, .52 for the high school record, and .58 for the three predictors combined. Thus, the validity coefficients of the tests, together with the high school record are only .06 greater than the high school record alone. [258:166] While noting that the SAT's contribution to the prediction of first-year grades largely depends on how the validity coefficients are interpreted, Slack and Porter stated that, in general, the SAT adds little to the prediction of college grades over the high school record alone. [166:167]

Although the Slack and Porter study employed different statistical techniques in its analysis of the predictive validity of the SAT, the Nairn study found similar results in its analysis of the Ford and Campos data. [258:167] Specifically, Nairn reported that, according to figures gathered from 827 different ETS validity studies conducted between 1964 and 1974, the SAT delivers a "percentage of perfect prediction" of 11.9 percent in predicting first-year grades. [188:60] Other ETS aptitude tests provide similar percentages of perfect prediction: 13 percent for the LSAT, 11 percent for the GRE, and eight percent for the GMAT. [188:61-62] Nairn said that in order to determine how often chance would predict grade ranking within a group as well as an ETS aptitude test, the tests' percentage of perfect prediction is subtracted from 100 percent. [188:64] Nairn's calculations show that the SAT scores, on the average, predict a candidate's "grade rank no more accurately than a pair of dice." [188:65] Furthermore, since SAT test scores are not the only factor in admissions decisions, Nairn asserted that their effective contribution to grade prediction is less than 12 percent. [188:65]

Additionally, Nairn said that data from the 1964-74 validity studies showed that including SAT scores in the prediction process improved the prediction of college grades by five percent or less. "This thin margin, the extent to which ETS scores improve the prediction already offered by previous grades," Nairn emphasized, "is the single thread, the single rational function, from which the ETS aptitude testing empire hangs." [188:66] Nairn stated that other forms of assessment such as biographical questionnaires, personal rating scales, persistence in staying in school, previous accomplishments, and other information have been found to predict first-year grades nearly as well or better than test scores. [188: 68, 72, 77]

TABLE 9.--1973-74 College Bound Seniors Classified by SAT Average and Family Income*

	Reported Family Income				
SAT Average	$ 0-5,999	$6,000-11,999	$12,000-17,999	$18,000+	Average income**
750-800	17	117	169	415	$24,124
700-749	239	1,172	1,752	3,252	21,980
650-699	686	3,994	5,683	9,284	21,292
600-649	1,626	9,352	12,187	17,992	20,330
550-599	3,119	17,042	20,822	28,151	19,481
500-549	4,983	26,132	29,751	37,400	18,824
450-499	6,663	33,209	35,193	41,412	18,122
400-449	8,054	34,302	33,574	37,213	17,387
350-399	8,973	29,762	25,724	26,175	16,182
300-349	9,622	21,342	14,867	13,896	14,355
250-299	7,980	10,286	5,240	4,212	11,428
200-249	1,638	1,436	521	325	8,639
Total Number	53,600	188,146	185,483	219,727	
Average SAT Score	403	447	469	485	

[*] The total number of students in this table (646,956) is very slightly smaller than the number (647,031) included in the analyses reported in *College Bound Seniors, 1973-74.* Students in this talbe must have had *both* SAT verbal and SAT mathematical scores and have reported family income on the Student Descriptive Questionnaire. Students with only one SAT score were included in *College Bound Seniors.*

[**] From *College Bound Seniors, 1973-74.*

SOURCE: *Test Scores and Family Income.* Princeton, NJ: Educational Testing Service, 1980, copyright 1980 by Educational Testing Service, 1980. p. 7. Used with permission.

According to Rex Jackson of ETS in an article appearing in the August 1980 *Harvard Educational Review*, Slack and Porter ignored the summary table in the Ford and Campos report and selected one sentence from the text which discussed only a portion of the data, which Slack and Porter then contrasted with their own findings. Jackson called the implication that the Ford and Campos results had been reported inaccurately "a serious injustice to the authors." [134.1:389]

In response to Nairn's charge that rolling dice is nearly as good as using test scores in the admissions process, ETS charged that Nairn's claim was based on faulty statistics. [77:16] ETS reported that Nairn used an incorrect value for the characteristic validity of the SAT (.345), which is the average of the separate validities of the verbal and mathematical parts of the SAT. Nairn should have used the greater predictive validity of the total test which is .41. After taking an erroneous validity of the SAT, Nairn squared the validity coefficients and multiplied the results by 100, thus arriving at the numbers 13 for the LSAT, 12 for the SAT, 11 for the GRE, and 8 for the GMAT. [77:16] Then, by taking the complements of these numbers, Nairn stated that "rolling dice will be as accurate as using test scores from 87 to 92 percent of the time." [77:16] ETS said it was not possible to know how Nairn reached this "misconception." [77:17] ETS emphasized that

> If the SAT were valid, it would *be* no *worse* than a random predictor, such as a pair of dice. If predictions based on an invalid test and random predictions were compared for a large group

TABLE 10.--Median Validity Coefficients of Combined-Sex Samples for the SAT-V, SAT-M, High School Record (HSR), and Combined Predictors, 1964-74

Source of Date	Year	Number of Colleges	SAT-V	SAT-M	HSR	Combined Predictors	Increase over HSR Alone
Ford & Campos (1977, Table 5, p. 9)	1964	62	.37	.32	.55	.62	.07[a]
	1965	80	.36	.29	.56	.61	.05
	1966	76	.33	.28	.53	.57	.04
	1967	62	.36	.28	.52	.57	.05
	1968	41	.40	.33	.56	.62	.06
	1969	45	.39	.33	.53	.59	.06
	1970	129	.37	.29	.49	.56	.07
	1971	32	.34	.28	.48	.53	.05
	1972	101	.35	.33	.50	.57	.07
	1973	89	.38	.36	.50	.56	.06
	1974	110	.42	.39	.50	.58	.08
Ford & Campos Summary[b] (1977, p. 3)	--	--	.40	.35	.50	.58	.08[c]
Slack & Porter's Calculations[d] from Ford & Campos (1977, Table 5, p. 9)	--	--	.37	.32	.52	.58	.06[e]

[a] These are our figures, calculated from Ford and Campos (1977) by subtracting the median validity coefficients for the high school record alone from the median validity coefficients for the combined predictors (combined-sex samples).

[b] These are the "averages" for the SAT-V, SAT-M, HSR, and combined predictors presented by ETS in the Ford and Campos summary (1977, p. 3).

[c] We calculated this increase over the HSR alone by subtracting the "average" for the HSR alone (.50) from the "average" for the combined predictors (.58) (Ford & Campos, 1977, p. 3).

[d] These are weighted means; we calculated them from the median validity coefficients reported in Table 5 of Ford and Campos (1977, p. 9). In the absence of information on the number of students in each sample, we assumed that the number of students was proportional to the number of colleges and used the number of colleges for each year as the weighting factor for that year's validities. For each variable, our calculations of the weighted means, means, and medians of the median validity coefficients (in that order) are as follows: SAT-V = .37, .37, .37; SAT-M = .32, .32, .32; HSR = .52, .52, .52; Combined Predictors = .58, .58, .57; increase over HSR = .06, .06, .05.

[e] We calculated this increase over the high school record alone by taking the difference between the weighted mean of the median validity coefficients of the high school records (.52) and the weighted mean of the median validity coefficients of the combined predictors (.58).

SOURCE: Slack, Warner V. and Douglas Porter. The Scholastic Aptitude Test: A Critical Appraisal," *Harvard Educational Review*, May 1980, p. 166. Copyright 1980 by President and Fellows of Harvard College. Used with permission.

> of students, the test would give predictions closer to students' obtained GPAs for about 50% of the cases, and random predictions would be better for the other 50%. Predictions based on valid information will be better than random predictions and will certainly more closely approximate obtained GPAs for more than 50% of the applicants. Yet, according to Nairn, dice will be as good as or better than the SAT, a test of proven validity, 88% of the time. Like citations of evidence in the Nairn report, these dice should be carefully inspected. (emphasis in the original) [77:17]

ETS further stated:

> In arriving at a figure of five percent, which in his chapter title he calls "five percent of nothing," Nairn switches statistics from r^2 to another index. In fact, he uses . . . the "coefficient of forecasting efficiency," which is equal to $1-\sqrt{1-r^2}$ The actual coefficients for the 1974 data (not reported by Nairn) are 13.4 for high school grades alone and 18.5 for tests and grades combined, which do indeed differ by 5.1. This index has a theoretical range from 0 (for random prediction) to 100 (for perfect prediction) and can be interpreted in percentage terms, though the interpretation is not a particularly simple one. Since "perfect prediction" cannot be attained, it is reasonable to consider the contribution of tests in relation to the level of prediction offered by grades alone. Simple arithmetic (5.1 divided by 13.4) indicates that combining test scores with grades increases the value of Nairn's chosen index *by* 38%, *not* 5%. (emphasis in the original) [77:19]

George Hanford of the College Board reported that Cameron Fincher's 1974 report on the use of the SAT in the university system of Georgia, which included 19 institutions over a 13-year period, found an average correlation of SAT scores with first-year college performance of .49, on a scale of 0.00 to 1.00 (perfect prediction). This study also showed that while high school grades are a better predictor with a correlation of .54, the combination of SAT scores and grades together increased the correlation to .65, "a strong incremental addition to the validity of the prediction." Hanford reiterated that the SAT is only one element in the admissions process, which often includes interviews, activities reports, and letters of recommendation. "The question is not whether the SAT is perfect," Hanford said, "but whether it is useful." [121:3]

Testing the tests.-- Because the consumers of education tend to accept unquestioningly the psychometrician's claims about aptitude tests, Nairn charged that ETS's claims about its tests constitute "a specialized variety of fraud." [188:58] This "fraud" appears to be respectable because it adheres to the rules of the psychometric profession and because political and economic institutions have been receptive to the test makers' verdicts about individuals' potential, Nairn stated. [188:58] Over the past 30 years, Ralph Nader estimated that the "pervasive power of ETS" has probably affected some 90 million people who have had their educational and occupational prospects shaped by its standarized tests. [184:1] Nairn contended that ETS should not be permitted *to define* consumer choices in terms of its tests nor should it "be given an intellectual monopoly to complement its economic one." [188:382] Further, he stated that reliance on multiple-choice tests unfairly restricts "the means by which individuals can demonstrate their skills and potential," [188:384] and that the tests can also hinder the development of essay writing and scientific problem-solving skills. [188:386] Nairn further stressed that the tests, when validated against first-year grades, "improve efficiency of prediction by an average of only three to five percent." [188:388] He also questioned the rationality of ranking applicants "with tests that provide only an eight to 15 percent improvement over blind chance in prediction of a short-term criterion such as first-year grades." [188:391] Nairn emphasized that standardized tests cannot purport to measure the variety of skills necessary for success in any academic or post-academic endeavor. [188:391]

Nairn charged that ETS has an overwhelming economic and political interest in maintaining the current testing system. In order to maintain this system and defend its existence, ETS has developed an extensive and subtle network of influence. This influence extends to the upper reaches of government, where ETS has had associations with four Commissioners of Education from 1962 to 1972.* It also extends to about 2,000 paid consultants at thousands of universities, colleges, educational associations, and local school districts across the country. [188:396] Nairn believes that reforming the test system must focus on ETS, the organization which develops and administers tests for thousands of institutions. Both Nader and Nairn asserted that they plan to take a leading role in the "national examination of the examiners." [188:384] Part of their upcoming agenda will feature the following considerations:

- "What are the effects of restricting the information system. . . to multiple-choice tests developed and chosen by a single organization?" [188:384]
- "What are academic and other institutions trying to accomplish with their admissions policies?" [188:387]
- "Precisely what goals and whose interests do those policies actually serve?" [188:387]
- "Should applicants have to bear the cost of the tests?" [188:389]
- "Is the ETS ranking really a meritocratic ranking?" [188:390]
- "What is the purpose of demanding that alternative admission standards rank candidates in a similar fashion?" [188:390]
- "Can significant reforms be implemented given ETS' current position?" [188:394]

ETS, on the other hand, asserted that it has worked diligently to help both candidates and institutions keep test scores in perspective. Although test scores are only one factor in the admissions process, many critics of testing have ascribed to them an importance that is substantially out of proportion, Turnbull said. [68:4] When testing critics publicize this fallacy, Turnbull added, they increase the likelihood of "test anxiety" among students, encourage others to seek expensive coaching schools, and instill an unfounded belief in high-scoring students that their test scores will earn them admission to the institution of their choice. [68:4]

Standardized tests, Turnbull emphasized, are the same for all candidates and reduce the unfairness that would result if college admissions officers reviewed only high school grades, which vary considerably among schools and teachers. [285:2] ETS stated, however, that personal qualities such as judgment, wisdom, creativity, idealism, stamina, and determination are highly valued and sought, but they are difficult to assess in other than subjective terms. [77:24]

---

*The Commissioners of Education included Francis Keppel, 1962-66, former president of the Carnegie Foundation for the Advancement of Teaching, a founder of ETS; Harold Howe II, 1966-68, a former College Board trustee, and vice-chairman of the College Board's Advisory Panel on the SAT score decline; James E. Allen, 1969-70, who served as a consultant to ETS; Sidney P. Marland, Jr., 1970-72, former president of the College Board and a trustee of ETS. Other prominent government officials with whom ETS has had associations include John Gardner, former president of the Carnegie Foundation, former ETS trustee, and former Secretary of Health, Education, and Welfare; Juanita Kreps, former ETS trustee and U.S. Secretary of Commerce; Willard W. Wirtz, former U.S. Secretary of Labor, and chairman of the College Board's Advisory Panel on the SAT score decline; former U.S. Congressman John Brademas, a former ETS trustee and former head of the House Subcommittee for Higher Education; Ernest Boyer, former U.S. Commissioner of Education, and president of the Carnegie Foundation; Shirley Hufstedler, former Secretary of Education, who served as a trustee of the Aspen Institute for Humanistic Studies, which has its international division located at ETS headquarters.

ETS said that it has no quarrel with Nader and Nairn proposing their views regarding criteria useful in making admissions decisions. But these views should not be considered the only correct values nor should they "be substituted for the considered judgment of the faculties and administrators of the educational institutions." [77:25] There are ample opportunities to contribute to thoughtful debate and improve testing in the admissions process, ETS observed; however, the Nairn report offered "no serious alternatives" and misrepresented many facts. [77:25]

The proper response to the campaign against testing and institutional autonomy, according to Hanford, ought to come from the professionals in guidance and admissions, who are involved in and are best informed about the programs and services under criticism. [121:5-6] Hanford accused Nader of "reckless and inflammatory attacks" which infer that the "purpose of these tests is to perpetuate an unjust social system by denying opportunities to those who rightly seek them." (emphasis in the original) [121:6] Furthermore, recently introduced legislation to regulate testing nationally, Hanford noted, could mean the demise of accountability programs and produce a ripple effect which could spread from admissions tests, to competency tests, measures of accountability, bar examinations, civil service examinations, and even driving tests. [123:9] Hanford reiterated that the SAT is not a perfectly accurate predictor, but when it is combined with high school grades, it provides the best predictor known. [123:8] "A nation that throws out its tests as Ralph Nader would have us do," Hanford cautioned, "is in danger of throwing out both its values and its standards." [123:9]

In an address to the Middle States Association of Colleges and Schools in December of 1979, Hanford posed these questions to his audience for consideration:

- "What *are* the viable alternatives to admissions tests?" (emphasis in the original) [123:10]
- "How much does your institution rely on tests, and test scores. . . in the guidance process if it's a school. . . or in the admissions process if it's a college?" [123:10]
- "Should we, and if so how could we. . . your schools and colleges as institutions. . . or the Middle States Association and the College Board as organizations. . . better respond to the concerns of the consumer movement? [123:10]
- "Should we *be*, and if so how can we *all* become more accountable, or perhaps, how can we bring about a better perception of our accountability?" (emphasis in the original) [123:10]
- "Can there be standards without assessments?" [123:10]
- "What is the minimum--where is the floor?" [123:10]

THE COACHING CONTROVERSY:
SPECIAL PREPARATION FOR THE SAT VS. NO SPECIAL PREPARATION

The highly publicized decline in SAT test scores is one reason for increased interest in the effects of coaching on the SAT and other standardized admissions examinations. [136:9] Many educators and much of the public believe that coaching can make the critical difference in gaining admission to college and graduate school. Scoring well on a standardized admissions examination is very important for those applying to a prestigious college or university. More than 90 percent of those admitted to Princeton University, Smith College, Stanford University, Wellesley College, Brown University, the University of Chicago, and the Massachusetts Institute of Technology score over 500 in *both* the SAT verbal and mathematical tests. [90:11]

In addition, Harvard requires students to score between 500 to 800 on each test; Yale requires students to score a minimum of 670 on the SAT-V and 680 on the SAT-M; both Pennsylvania and Columbia Universities require scores of 650 on the SAT-V and 660 on the SAT-M; Pennsylvania State has an acceptable range of scores between 450-600 on each test; Emory requires a 550 on the SAT-V and 600 on the SAT-M; Rutgers 490 on the SAT-V and 540 on the SAT-M; and George Washington University, a combined total of at least 1,000.* [194:35] The 1980 SAT score averages for college-bound seniors were 424 on the SAT-V and 466 on the SAT-M, a total of 890 points out of a maximum of 1,600. [6.2:5]

According to test critic Allan Nairn, it has been estimated that as many as 40 percent of the nation's public colleges and 20 percent of private colleges use test scores as automatic cutoff points. [35:190] As more individuals seek coaching to improve their test scores and their chances of admission to select colleges and universities, the debate over the effectiveness of coaching for the SAT and other tests will continue to escalate.

The College Entrance Examination Board, on the other hand, maintains that interest in special preparation or coaching stems from an exaggerated view of gaining admission to college at the undergraduate level. Test scores are only one component in the admissions process along with high school transcripts, letters of recommendation, previous experiences, and activities. [50:1] Traditionally, the CEEB has claimed that the SAT measures abilities developed over a lifetime. As a result, the College Board believes that the SAT is not susceptible to *short-term* instruction, special preparation, or coaching. [122: 1, 3] The coaching issue is complex because the term "coaching" can refer to several types of test preparation. [66:1] Added to the complexity in understanding the effects of coaching is that students' test scores change over time due to growth in verbal and mathematical reasoning abilities and "a lack of perfect precision in the measuring instrument." The College Board has estimated that students' verbal and mathematical scores will increase approximately 15 to 20 points each between the spring of a student's junior year and the winter of his or her senior year. Additionally, about five percent of all students will notice a score increase of approximately 100 points or more, while one percent will experience a score decrease of a similar amount regardless of whether special preparation for the SAT occurs. [50:2] Earlier studies of coaching or special preparation that included control groups, according to ETS, resulted in test score gains of less than 10 points for the verbal portion of the test and less than 15 points for the mathematical portion of the test. [177:3]

---

*In 1980 the College Board reported that only 5.5 percent of all SAT takers or about 55,000 students had a SAT average of 500 or higher and parents who could contribute $5,000 or more to their college education. [6.2:8]

In a 1965 statement that was essentially unchanged until 1979, the College Board said: "Intensive drill for the SAT, either on its verbal or its mathematical part, is at best likely to yield insignificant increases in scores." [50:1] But in light of two reports released by the Federal Trade Commission and growing criticism directed toward standardized testing in general, test agencies have begun to make subtle and potentially important adjustments. In the past the College Board applied the term "coaching" only to intensive drill or "cramming" on sample test questions. [50:2] Yet recent statements from the College Board point out that cramming is not the same as "formal instruction in reading, comprehension and verbal and mathematical concepts," and *short-term drill* is likely to have little effect; well designed, *longer preparation* can have greater effect." (emphasis in the original) [122:3-4] George Hanford, president of the College Board, observed that "when something stops being short-term cramming and becomes long-term tutoring, there is a distinction." [168:E11] Although largely a matter of semantics, this fine distinction, according to one education writer, may eventually lead to an acknowledgment by testing agencies that coaching is effective in some cases. [136:9] Critics of standardized tests go further, and claim that if coaching improves test scores substantially, then the SAT does not measure one's aptitude or intelligence. [188:x]

Commercial Coaching Schools

The test-coaching industry is growing at an extraordinary pace. There are few state licensing restrictions and almost no entry barriers to this industry. [90:36] In 1976 commercial coaching schools spent $400,000 on advertising, and nationwide sales yielded approximately $10 million. [90:31] According to *Forbes* magazine, commercial coaching in 1979 grossed $60 million. [293:21] Approximately 50,000 students enrolled in coaching schools in 1976. [90:31] Tuition charges range from a minimum of $40 to a maximum of $500. [90:31-32; 194:26] The nation's largest commercial coaching enterprise, the Stanley H. Kaplan Educational Center (SHK), began over 30 years ago, enrolls more than 30,000 students annually, and its 1978 sales reached $9 million. [168:E11; 237:1, 20]

Commercial coaching schools vary considerably in their curricula, instructional methods, and course materials. The coaching curricula may include teaching test-wiseness and exam taking techniques, short-term instruction which includes substantive content, or intermediate-term instruction. At the Columbia Test Preparation Institute students review analogy questions and are taught to recognize "reverse meaning," "word association," "ungrammatical answers," and to avoid inappropriate answers. In order to dispel student anxiety, another school, the Sexton Center, offers remedial mathematics instruction in addition to its regular program. Other coaching enterprises such as the Guidance Center in Santa Monica, California, and the Dulac School in New York City provides "individualized counseling and psychological reassurance." [237:1, 20]

Coaching schools utilize written materials, commercially available review books, and self-designed review books and materials. SHK also has a tape library which offers students about 200 hours of additional review. [90:32] The Heights Study Center of Washington, D.C., offers students 20 hours of classes equally divided into verbal and math sessions with 50 hours of homework. Students who maintain a perfect homework-attendance record but fail to add 100 points to their previous combined verbal-math score are entitled to a 70 percent refund or a full tuition scholarship at a subsequent session. [124:2] At the larger preparatory schools, staffs of researchers are constantly updating and writing practice questions. Several coaching schools, however, apparently pay individuals to take standardized examinations for the purpose of remembering the test questions. [90:32-33] Upon learning that some questions appearing in a 1978 edition of the SAT were identical to questions distributed in a private coaching course, ETS filed a copyright suit and was awarded $15,000 in damages against Tuchman Tutoring Service of Hartsdale, New York. Students who had access to the test questions had their scores cancelled. [276:2]

Boston Regional Office Report on the Effects of Coaching on Standardized Admission Examinations

On October 13, 1976, the Federal Trade Commission (FTC) directed its Boston Regional Office (BRO) to conduct an investigation to determine if test preparers, review courses, coaching schools, persons, or partnerships, were engaged in unfair or deceptive practices affecting commerce in violation of Section 5 of the Federal Trade Commission Act. [90:5] Arthur E. Levine, BRO staff attorney, headed the investigation to determine whether test preparation centers had a reasonable basis for their claims that scores on standardized tests could be increased to the amount advertized. [90:6] Some examples of claims made by coaching enterprises were:

- "The scores of many of our students have jumped 250 points or more." [90:171]
- "The average increase of students who take our course is 80 points with some improving as much as 150 points." [90:173]
- "...If an average improvement of 30 points results merely from having taken the test before, a still greater score advantage would be expected to follow an effective instructional program...." [90:173]

The BRO study was one of the largest coaching studies ever undertaken and one of the few studies devoted to investigating commercial coaching enterprises. Since there were problems in identifying defunct coaching schools, inadequate record retention by the schools, and fewer numbers of students enrolled in preparation courses prior to 1974, BRO's investigation was limited from October 1974 through December 1976. Nonetheless, the period covered offered BRO's staff the large sample sizes necessary to conduct a statistical inquiry. [90:26, 28] The sample consisted of 2,741 students, which included 1,738 who were uncoached and 1,003 who were coached. [113:1] The investigation was limited to the Scholastic Aptitude Test and the Law School Admission Test. [90:27] With regard to the SAT, two commercial coaching enterprises were studied, the Stanley H. Kaplan Educational Center (called School A) and the Test Preparation Center (called School B). [113:1] SAT coverage was confined to the New York area, but the LSAT coverage included coaching schools in Boston, Chicago, Los Angeles, and New York. [90:50]

The Boston Regional Office observed that there were two methods for determining the effects of coaching upon SAT and LSAT scores:

(1) conducting a controlled experiment and (2) conducting an experiment using existing conditions. [90:46] Although an experimental approach was preferable, the BRO felt that it would be expensive, time-consuming, and require denying students access to commercial coaching. [90:48] Consequently, the BRO study employed a nonexperimental design: students being examined were not randomly assigned beforehand to either a coached or noncoached group. [113:1]

A baseline was established to compare coached students against noncoached students. PSAT scores provided the SAT baseline, while undergraduate GPAs provided the LSAT baseline. These baselines, according to BRO were selected because of their availability, convenience, and reasonableness. From these baselines, expected average test scores were generated for both coached and uncoached students. Comparisons of these average scores were then used to analyze the practical effects of commercial coaching. [90:59] BRO, however, cautioned that the uncoached, or control groups, may not have been completely gleaned of all coached students because some students may have attended an unidentified coaching school, may have experienced intensive self-preparation, or may have been coached by the nonprofit portion of the coaching industry. [90:60-61]

BRO's statistical analyses for SAT takers who were coached in the period intervening between the PSAT and the SAT, and who took the SAT only once, indicated that:

- Average verbal SAT increases attributable to School A's coaching ranged from 40 to 76 points above the uncoached group. Students who scored higher on the PSAT received the greater benefits.
- Average math SAT increases in School A were 40 points above increases of the uncoached group over the entire range of the PSAT.
- Students who were coached by School A prior to the SAT showed a combined net increase between 80 and 116 points above a matched control group.
- Although somewhat helpful, School B's coaching appeared to be less effective than School A's for first-time takers.
- Compared to the uncoached group, average verbal SAT scores attributed to School B's coaching ranged from an increase of 60 points for low-PSAT students to a loss of 33 points for high-PSAT students.
- Average math SAT scores ranged from an increase of eight points for low-PSAT students to an increase of 40 points for high-PSAT students above increases recorded by uncoached students.
- Students who were coached by School B prior to the SAT had scores ranging from -25 to +100 points compared to the uncoached group. [90:62,63]

Students who took the SAT twice, yet did not receive coaching prior to the examinations, showed improved verbal and math scores proportional to PSAT scores. Two possible reasons were suggested by BRO for the pattern of improvement: (1) "persons with higher initial test scores tend to continue to develop further than those with initially lower test scores" and (2) "persons with higher test scores find the SAT somewhat 'learnable' even without coaching. [90:70]

Comparisons of students who took the SAT twice and who were coached *before* the first examination revealed that School A's coaching was effective for all students and that School B's coaching was effective for low-PSAT students, but perhaps counterproductive for high-PSAT students. It was observed, however, that School B's high-PSAT students regained most of their lost ground on their second SAT attempt. [90:74] Other findings related to two-time takers of the SAT were:

- Average verbal SAT score increases for students from School A, when compared to scores for a matched control group, ranged from 56.7 points to 32.7 points from the low- to the high-PSAT scores, respectively. [90:74]

- Average math SAT score increases ranged from 57 points for low-PSAT students to 16 points for high-PSAT students. A comparison of second SAT scores showed an average difference of 74 points for low-PSAT students, compared to almost no difference at all for high-PSAT students. [90:79]
- Average verbal SAT increases attributable to School B's coaching were 34.4 SAT points at the low end of the PSAT scale, but showed losses of 55.6 SAT points at the high end of the PSAT scale. On their second SAT, School B's students showed losses of 22.8 SAT verbal points at the high end of the PSAT scale, compared to a matched control group. [90:74]
- Average math SAT scores showed improvements ranging from 42 points at the low end of the PSAT scale to losses of 54 points at the high end of the PSAT scale. On their second SAT, math scores increased 77.6 points from their first SAT math score at the high end of the PSAT scale, compared to an increase of 36.6 SAT math points for the control group. [90:79]
- School B's initial coaching helped low-PSAT students but higher-PSAT students made greater improvements than uncoached students on their second SAT. [90:79]

Comparisons of students who took the SAT twice and who were coached *between* examinations revealed that the Kaplan Center's coaching was effective and the Columbia Test Preparation Institute's coaching provided "minimal benefits":

- School A's coaching showed increases over the entire PSAT range with SAT second verbal scores increasing from 39 to 43 points. [90:84]
- School A showed second math increases between 22 and 76 points from the low to the high range of PSAT scores, while the control group showed increases of one to 35 points. School A's students showed a net gain for second SAT math scores between 21 and 41 points. [90:89]
- School B's coaching showed increases of 1 to 7 verbal points over the entire PSAT range. [90:84]

According to BRO, comparisons of results for those students who were coached before the first SAT and those who were coached between the first and second examinations indicated that coaching and SAT experience offered different benefits: (1) coaching without prior SAT experience appeared to help low-PSAT students; (2) examination experience with or without coaching appeared to help high-PSAT students; and (3) coaching after an examination appeared to help everyone. [90:89]

The results of BRO's statistical analyses indicated that coaching was highly effective for the SAT. But the results for the LSAT showed only marginal effects due to coaching. [90:157] Yet according to Allan Nairn, the SAT coaching courses studied raised scores by an average of more than 100 points, while one of the two LSAT courses studied showed scores improved by an average of 60 points. [188:102]

In its summary, BRO stated that standardized examinations used in admissions decisions were susceptible to short-term preparation or coaching found in numerous commercial coaching enterprises. The analyses indicated that there was a "statistically significant difference between the score increases obtained by coached and uncoached individuals." It was noted that the score increases varied somewhat both between and within the examinations. Nonetheless, the BRO stated that the score increases due to coaching could determine who is admitted to undergraduate and graduate institutions.

Since the availability of coaching was directly related to the ability of individuals to pay tuition at coaching schools, BRO concluded that "coachable standarized admission examinations create financial barriers to educational opportunities in direct conflict" with the Congressional declaration that it is "the policy of the United States of America that every citizen is entitled to an education to meet his or her full potential without financial barriers." [90:1-2] Furthermore, BRO said that standarized admission examinations discriminate against any individual who either "cannot afford the cost of commercial preparation or elects not to attend a commercial preparation course even if he can afford it because of acceptance of the dogma promulgated by the test makers, test administrators, and test users over the past twenty years that coaching is valueless." [90:3] Moreover, the existence of a single coaching school (School A) that could actually increase an examinee's scores on standardized admission examinations, BRO stated, indicated imperfections in the reliability and validity of tests such as the SAT and LSAT. [90:2]

The BRO report recommended the following actions to eliminate the unfair practices it claimed to have found: (1) publication of a report to Congress; (2) initiation of formal investigations of the Educational Testing Service, College Entrance Examination Board, and the American Bar Association to determine the extent of the FTC's jurisdiction over them; (3) supplementation of the statistical analysis in "an attempt to create uncoachable, more valid and reliable admissions examinations;" (4) initiation of litigation against any coaching school which seriously violates Section 5 of the FTC Act. [90:4-5]

In September 1978, BRO's study titled *The Effects of Coaching on Standardized Admission Examinations* was submitted to the U.S. Bureau of Consumer Protection. A month later, the FTC decided to postpone release of the study in order to reanalyze its data. After reviewing the methodology of the BRO study, the Bureau of Consumer Protection stated that more sophisticated analyses of the survey data were necessary in order to draw "reasonable conclusions." [112:2] According to one education writer, the delay in releasing the BRO study fueled speculation that the FTC was sitting on the report because its conclusions favored the coaching industry. [167:A10] In January 1979, Arthur Levine, the attorney in charge of the BRO study, resigned from the FTC because of the commission's "hesitancy to release the nationwide study." Amidst the speculation, ETS president Turnbull strongly denied that ETS had influenced the FTC to delay releasing the BRO report. [171:24]

The Bureau of Consumer Protection's Revised Statistical Analyses of the BRO Report

Some six months after the conclusion of the Boston Regional Office's report, the Federal Trade Commission's Bureau of Consumer Protection (BCP) completed its "official" revised statistical analyses of the impact of coaching on SAT scores. BRO's study served as the basis of the BCP's reanalyses, and, like its predecessor, the *Revised Statistical Analyses* (March 1979) employed a nonexperimental design. However, unlike the BRO study, the BCP's reanalysis stated that there was "no reason to expect the coached and uncoached groups to have similar demographic and personal characteristics." Consequently, if these differences were uncontrolled in the statistical analyses, then the differences between the coached and uncoached groups might not be attributed solely to coaching. [89:6]

The BCP's reanalyses showed that a comparison of demographic profiles of the coached and uncoached students indicated considerable differences between the two groups. Nearly 31 percent of the coached students ranked in the top 10 percent of their class, compared to 21 percent of the uncoached students; 41.2 percent of the coached students reported family income of $30,000 or more, compared to 17.2 percent of the uncoached students; 44.7 percent of the coached students attended private or parochial high schools, compared to 24.6 percent of the uncoached students; 55.3 percent of the coached studients reported their latest grade in English was an "A," compared to 35.5 percent of the uncoached students; and 48.3 percent of the coached students reported their most recent grade in mathematics was an "A," compared to 29.6 percent of the uncoached students. [89:8-10]

Even with these personal and demographic differences among coached and uncoached students, the *Revised Statistical Analyses* found that coaching was effective at one of the two SAT-coaching schools studied (School A), and that students increased their SAT scores on the average of 25 points each on the verbal and mathematical examinations. Moreover, students who attended the most effective coaching school appeared to be underachievers on standardized examinations, that is, they scored lower on these tests than would have been expected according to personal and demographic characteristics including grades in school and class rank. [89:i]

Table 11 presents data on the overall impact of coaching as reported by BCP. Using data for first-time SAT takers pooled across all the test dates (subsample 5), BCP found that coaching at School A provided an average increase of 29.7 points to students' verbal SAT scores and 19.2 points to their math SAT scores. In contrast, the results for School B provided -1.8 points for the verbal SAT and +5.4 points for the math SAT. The study observed that the impact of coaching at School A was nearly 30 points on the verbal and 20 points on the mathematical section compared to almost "zero points" at School B. [89:21]

In order to analyze the impact of coaching upon SAT scores and take into consideration the effects of self-selection, BCP conducted a regression analysis to estimate the number of test points that could be ascribed to the coaching at both schools. [89:28] As shown in Table 12, the first SAT pooled across all the exam dates that considered the self-selection factor indicated that the effects of coaching on both the verbal and math SAT exams were considerably less than those reported in Table 11. The mean number of points attributable to coaching at School A on the verbal SAT (subsample 5), was 11.5 points compared to 29.7 points without the adjustment for self-selection, and the mean number of points on the math exam was 5.5 points compared to 19.2 points. The results for the second SAT pooled across the test dates were different. For School A the number of points ascribed to coaching were less than those reported earlier (16.2 and 16.6 compared to 27.2 and 28.4 on the verbal and math SAT's, respectively). These results were still statistically significant, according to BCP, that is, after the adjustment for self-selection, School A had a beneficial influence on both examinations. The results for School B were less than those reported in Table 11, and were not considered significant. [89:30]

An analysis of the results for each of the first four subsamples representing the four different test dates revealed mixed results for School A. For the verbal SAT exam, coaching had a statistically significant effect for only two of the four test periods. For the math portion of the SAT, School A was effective only for one of the four exams. BCP stated that "if one believes that coaching must be consistently effective before one can conclude that coaching schools work, one can not state that coaching at School A works (at least after taking into account self-selection)." [89:30-31] Both before and after taking the effects of self-selection into account, the results indicated that coaching at

TABLE 11.--Overall Impact of Coaching: Mean Number of Points (and Confidence Interval) Contributed by Coaching School

	School A Stanley H. Kaplan Education Center		School B Test Preparation Center	
Subsample*	Verbal SAT	Math SAT	Verbal SAT	Math SAT
1. 1st SAT 4/75 (n = 476)	18.0 (2 - 34)**	17.1 (0 - 34)	***	***
2. 1st SAT 4/76 (n = 658)	44.5 (33 - 57)	26.5 (14 - 39)	3.5 (-10 to 17)	.2 (-15 to 15)
3. 2nd SAT 11/75 (n = 359)	26.9 (11 - 43)	22.4 (5 - 40)	0.5 (-25 to 26)	27.9 (0 - 56)
4. 2nd SAT 11/76 (n = 438)	25.4 (14 - 36)	30.7 (19 - 42)	9.5 (-11 to 30)	2.0 (-20 to 24)
5. 1st SAT - Pooled Time Periods (n = 1578)	29.7 (21 - 39)	19.2 (10 - 28)	-1.8 (-13 to 10)	5.4 (- 6 to 17)
6. 2nd SAT - Pooled Time Periods (n = 1176)	27.2 (19 - 35)	28.4 (20 - 37)	5.5 (- 9 to 20)	3.0 (-12 to 18)

*The sample sizes reported here vary from those reported in the methodology section of the text because some students did not respond to one or more items on the Student Demographic Questionnaire and were therefore dropped from this analysis.

**Confidence intervals.

***No one taking this SAT exam received coaching at School B.

SOURCE: Federal Trade Commission. *Effects of Coaching on Standardized Admission Examinations: Revised Statistical Analyses of Data Gathered by Boston Regional Office of the Federal Trade Commission.* Washington, D.C.: FTC Bureau of Consumer Protection, March 1979, p. 20.

School B did not have a statistically significant effect. [89:31]

While BCP found some evidence that students who were not underachievers on standardized tests could benefit from coaching at School A, these results were not conclusive since data were available only for one test. [89:ii] Overall, the BCP study pointed out that before concluding that School A was effective and School B was not, additional analyses would have to be conducted. The BCP felt that it was plausible that self-selection itself was responsible for the effectiveness of coaching rather than the coaching per se. [89:22-23] Nonetheless, BCP stated in its *Revised Statistical Analyses* that "Even though is can't be firmly concluded that coaching will work for everyone, the results of the study do show that coaching can be effective for those who do not score well on standardized tests." BCP cautioned, however, that "if large numbers of students were to obtain coaching because they felt it was effective, they might be very disappointed if in fact coaching really is only effective for underachievers." [89:35]

TABLE 12.--Impact of Coaching Adjusting for Self-Selection: Mean Number of Points (and Confidence Interval) Contributed by Coaching School

	School A Stanley H. Kaplan Education Center		School B Test Preparation Center	
Subsample	Verbal SAT	Math SAT	Verbal SAT	Math SAT
1. 1st SAT 4/75 (n = 476)	-6.1 (-32 to 20)*	-1.4 (-27 to 24)*	**	**
2. 1st SAT 4/76 (n = 658)	26.8 (8 - 45)	7.5 (-11 to 26)	.8 (-20 to 22)	8.8 (-13 to 30)
3. 2nd SAT 11/75 (n = 359)	27.9 (1 - 55)	6.7 (-20 to 34)	8.3 (-36 to 53)	32.1 (-11 to 75)
4. 2nd SAT 11/76 (n = 438)	10.3 (-9 to 30)	19.4 (0 - 39)	-1.5 (-38 to 35)	3.2 (-34 to 41)
5. 1st SAT - Pooled Time Periods (n = 1578)	11.5 (-2 to 25)	5.5 (-8 to 19)	-2.3 (-20 to 15)	13.4 (-4 to 41)
6. 2nd SAT - Pooled Time Periods (n = 1176)	16.2 (2 - 31)	16.6 (2 - 31)	-3.9 (-29 to 21)	-4.4 (-29 to 21)

*Confidence intervals.

**No one taking this SAT exam received coaching at School B.

SOURCE: Federal Trade Commission. *Effects of Coaching on Standardized Admission Examinations: Revised Statistical Analyses of Data Gathered by Boston Regional Office of the Federal Trade Commission.* Washington, D.C.: FTC Bureau of Consumer Protection, March 1979, p. 29.

Criticisms of the BRO Report

In May 1979, the FTC released both the initial Boston Regional Office study and the "official" Bureau of Consumer Protection study. Unlike BCP's study, the BRO study stated that coaching was effective for the SAT and could be the determining factor in deciding who is admitted to college. The BRO study also stated that coachable standardized admission examinations created financial barriers to those pursuing educational opportunities and that, therefore, the tests were in direct conflict with the nation's education policy. The BRO study further recommended an investigation of the College Board, the Educational Testing Service, and the American Bar Association. [90:1-5; 114:1] While the BRO study found that students could improve their combined test scores by 100 points, the BCP's revised statistical analyses found that students could improve their test scores by 25 points each on the verbal and mathematical sections of the SAT. [188:102; 89:i]

BCP attached a one page disclaimer to the BRO study, explaining what it termed were "major flaws in the data analysis" which made the results unreliable. Albert Kramer, director of BCP, said that "we regard the study as being of limited significance." Among the flaws cited by

BCP were the following: (1) the BRO study failed to control for differences which existed in personal and demographic characteristics of coached and uncoached students; (2) without controlling for grades, it was not possible to know if SAT score differences between two groups were due to coaching or differences in skill represented by past school performance; (3) the report did not provide tests of statistical significance necessary to interpret the results; (4) findings were given for students who had a grade point average of 1.0 (a D average), and because there were no such students in the sample, it was not justifiable to predict the action of that group of students; and (5) the study's conclusions were limited due to the lack of controlled experimental conditions. [90] In releasing both studies Kramer said that the FTC wanted to stimulate further research and public discussion on the usefulness of coaching, but it was up to parents and students to determine whether commercial coaching is worth the expenditure of several hundred dollars. [95:3; 167:A10]

In response to BCP's disclaimer, Arthur Levine, former BRO staff attorney in charge of the initial investigation of coaching schools, said it was disturbing that after three and a half years and $300,000 the FTC had decided not to remedy a "glaring inequity in our national education system." [165:1] "I see this as an abrogation of their responsibility to consumers," he said. [167:A10] In a similar vein, Terry Herndon, executive director of the NEA, denounced the FTC for an incomplete job of investigating private coaching schools. Herndon said that it seemed the FTC was "aligned" with ETS and CEEB and, therefore, was unable to conduct an impartial investigation. [195:2] "People who can afford the cost of [coaching] schools," Herndon said, "thus have an unfair advantage at important transition points." [168:E11] According to testing critic Steve Solomon, the BRO study posed a political dilemma for the FTC. Solomon reported that an internal May 15, 1979, memo from Albert Kramer to FTC Commissioner Michael Pertschuk stated that: "The greatest consumer injury suffered by an applicant taking a standardized test comes not at the top or in the middle range, but as the student's score approaches whatever arbitrary cut-off the college or university imposes . . . .If the school's cut-off is 500, the difference between 520 and 495 is critical. Therefore, perhaps even the small coaching-attributed score gains acknowledged by ETS/CEEB portend some degree of unfairness to test-takers and would render deceptive their admonition not to be coached." According to Solomon, the FTC decided not to bring a deceptive trade practice suit against ETS because of its limited resources, political criticism of the scope of its activities, and the uncertainty of whether a nonprofit corporation fell under the FTC's jurisdiction. [260:4]

ETS, on the other hand, criticized the BRO report, stating that previous studies by ETS and others over a span of 26 years did not reach conclusions that would support the size of BRO's findings. Among the criticisms leveled by ETS against the BRO study were the following: (1) poorly designed research and erroneous conclusions could adversely affect students; (2) an apparent endorsement of coaching may encourage students to look for short cuts in preparing for standardized examinations; (3) students may be victimized by commercial coaching schools which make unsubstantiated claims and charge exorbitant fees; and (4) the studies did not take into account the provisions that ETS has already made regarding the preparation of students, including distribution of a sample test and discussion questions to test takers. Other criticisms of the BRO study included the lack of control for student motivation, disregard for general intellectual development over time, and the error present in all types of measurement. ETS further noted that the BRO study found no reliable results concerning the LSAT. [74:1-3] Former College Board President Robert J. Kingston said that the FTC reports could not "stand against the wealth of carefully directed research in this area," that shows the SAT is "highly resistant to coaching." Kingston further said that some individuals could gain from focusing their intellectual and emotional energies, but researchers still needed to know a lot more about the particular characteristics of students who enroll in coaching schools, as well as the aspects of coaching programs that seem beneficial to students. [113:2]

ETS's Reanalysis of SAT Data in the 1979 Federal Trade Commission Report

In May 1980, ETS held a press briefing in Washington, D.C. and presented a summary of its reanalysis of the FTC study conducted in 1979 by the Bureau of Consumer Protection. In "The Effectiveness of Coaching for the SAT: Review and Reanalysis of Research from the Fifties to the FTC," ETS stated that without random assignment of students, such as in the FTC study, one could not expect the effects of special preparation to be separate from a student's personal or background characteristics. Random assignment assures that there is "no systematic difference between the two groups," if proper controls exist, except that one group receives coaching and the other does not. Since the FTC study lacked randomization, there was no way to interpret the findings. [177:1] Such factors responsible for students seeking coaching may include lower than expected PSAT and SAT scores, in light of previous grades, or an attempt by the student to acquire high test scores to make up for a "prosaic high school performance." ETS's reanalysis observed that "the effects of self-selection are confounded with effects of the coaching treatment in nonrandomized studies and, consequently, self-selection factors afford plausible rival explanations for the results, or for part of the results, that might otherwise be identified as coaching effects." [177:2]

ETS further stated that an historical overview of the effects of coaching on the SAT prior to the FTC study showed that some studies utilized nonrandomized designs or had poor controls or small samples. [177:2] But the average effect of coaching in these earlier studies with control groups found score increases of less than 10 points for the SAT-V and less than 15 points for the SAT-M. ETS cited the Evans and Pike study (1973) which showed a 16 point increase on the SAT-M for students who were coached over seven three-hour sessions and who were assigned 21 hours of homework. For the SAT-V, ETS cited the Alderman and Powers study (1979) which found an eight point increase on the SAT-V for students receiving special preparation at eight secondary schools. ETS noted that "studies lacking control groups produced larger score effects but since they differed from control-group studies not only in design characteristics but in critical program characteristics, their interpretation was especially problematic." [177:3]

ETS included two specific reanalyses of the FTC study. The first covered three commercial coaching schools, including one which the FTC did not analyze because of its small sample of students. The overall results, however, were comparable to the FTC's findings, saying that "inconsistent and negligible effects for students at two schools and for students at the third school combined coaching and self-selection effects of about 20 to 35 points for both Verbal and Math scores." [177:4-5] Moreover, ETS found that a small sample of 13 black students attending one of the coaching schools that showed negligible effects overall exhibited larger coaching/self-selection effects than nonblacks on the SAT-V, but not the SAT-M. Also, students reporting low family income showed greater coaching/self-selection effects than those reporting high family income. [177:5]

ETS's second reanalysis took into account the growth rate in SAT scores over time, but did not adjust for self-selection factors or other factors not related to differential growth. [177:5] This reanalysis was performed on data from the largest coaching school which, according to the FTC, also reported a relationship between coaching school attendance and SAT performance. According to ETS, after taking into account the differential growth, the combined coaching/self-selection effect yielded about 11 points for the SAT-V and 30 points for the SAT-M. [177:6]

Overall, the summary of ETS's reanalysis of the FTC data revealed "considerable variability in the coaching/self-selection effects associated with coaching-school attendance, with negligible effects at two schools and an estimated combined effect for students at one school of about 20 to 30 points on SAT-Math and very likely about half that (as reflected in corrections for differential group growth rates) on SAT-Verbal. [177:6]

The ETS summary emphasized that "in any event, whether due to coaching or to personal and background factors, score increases of 20 to 30 points on a 200- to 800 point scale correspond to approximately three additional items correct, and the same effect might be obtained if students were to devote comparable time and effort to independent study or regular academic courses." [177:7] ETS further noted that no study has yet addressed the question of whether score increases reflect stable, long-term improvements in a student's verbal or mathematical reasoning skills or increases in testwiseness or what kind of instruction may be most beneficial for different kinds of students over a certain period of time. [177:8] The ETS summary recommended that test takers should be provided with test familiarization materials and other aids useful to test taking, prepared and distributed by the College Board. "Issues of equity of access to coaching programs that help overcome such extraneous test difficulty," the summary concluded, "become important to the extent that student differences in test-taking skills per se substantially influence test scores." [177:9]

ETS is planning to publish its complete reanalysis of the FTC data shortly. [122:1]

NEA's Analysis of SAT Data in the 1979 Federal Trade Commission Report

A July 1980 publication, *Measurement and Testing: An NEA Perspective*, reported the findings of the National Education Association's own analysis of the FTC data for students who had taken the PSAT once and the SAT twice. Using the FTC data base, the coached students in the sample were selected from the coaching school previously identified by the FTC as having "produced the *best* results." (emphasis in the original) The sample of 1,324 students included 625 who were administered the tests in 1975, and 699 students who were administered the tests in 1976. Students were assigned into groups based on whether they had been coached and when they were coached. In 1975, 65 students were coached between the PSAT and their first SAT; 105 students were coached between their first and second SAT; and 405 students were uncoached. In 1976, 118 students were coached between the PSAT and their first SAT; 173 students were coached between their first and second SAT; and 408 students were uncoached. Each student was used as his or her own control and no comparisons were made among the groups. The measure of growth employed in the study was the average increase between the PSAT and the second SAT. [194:24]

According to NEA, there were significant differences for each of the three subgroups in 1975 and 1976 between the PSAT and second SAT scores. (See Table 13.) This analysis of 1975 data for the group coached between the PSAT and the first SAT demonstrated an average increase of 114 points, while the group coached between the first and second administration of the SAT demonstrated an average increase of 104 points. In contrast, the uncoached group experienced an average gain of 44 points. An analysis of 1976 data for the group coached between the PSAT and the first SAT demonstrated an average increase of 143 points, the greatest average score gain reported in any coaching study. The group coached between the first and second SAT demonstrated an average increase of 135 points. The uncoached group experienced a gain of 60 points. [194:25]

NEA observed that coached students could increase their combined verbal and math SAT scores the second time they took the test more than uncoached students. [199:5] The analysis also noted that there appeared to be a relationship between average family income and coaching. [194:25] "Common sense would suggest that if one is taught four hours a week for ten weeks," the report emphasized, "there should be on the average some positive results." [194:26] The analysis further stated that NEA believed that differences of 25 or 50 points on the College Board's scale of 200-to-800 could make the difference in determining whether a student is admitted to the college of his or her choice. Since many families cannot afford to send their children to expensive coaching schools, the NEA called for all students to be coached free of charge. [194:26]

Slack and Porter's Criticisms of ETS

Warner V. Slack and Douglas Porter, of the Harvard Medical School, contended that the issue of special preparation is important because if students can raise their test scores by studying for the SAT, then the tests do not comply with Ryans and Frederiksen's 1951 definition of an aptitude test as "a device for measuring the capacity or potentiality of an individual for a particular kind of behavior." [258:155] Slack and Porter critically reviewed 29 studies on the effects of coaching. All of the studies were published prior to the release of the College Board's 1968 booklet *Effects of Coaching on Scholastic Aptitude Test Scores*, in which the College Board stated that "the evidence collected leads us to conclude that intensive drill for the SAT, either on its verbal or its mathematical part, is at best likely to yield insignificant increases in scores." [258:156, 161]

The authors noted that the College Board and ETS made repeated references to seven studies (Dear, 1958; Dyer, 1953; Frankel, 1960; French, 1955; French and Dear, 1959; Lass, n.d.; and Whitla, 1962) as the basis for their statement that special preparation is ineffective. [258:156] According to Slack and Porter, however, the best methods of special preparation were not employed in these studies. [258:158] They further charged ETS with failing to refer to other studies in the literature which show that coaching can be effective. They claim that ETS made no reference in its 1968 publication for students to two important studies, published in 1961 and 1965, which indicate that special preparation can improve students' scores under certain conditions, although Marron's 1965 study was, in part, supported by a grant from the College Board. [258:159; 202:1]

In particular, they charged ETS with ignoring a 1961 study by Nathaniel Pallone, a guidance counselor in a private college preparatory boys school. After analyzing SAT-V items supplied by ETS, Pallone designed both short-term and long-term courses to improve his students' test scores. The short-term preparation included a 90-minute daily training in verbal analogies,

TABLE 13.--Select Average SAT Point Gain Scores for Coached and Noncoached PSAT and Two-Time SAT Takers for 1975 and 1976

Year		Total Average Point Gain	Verbal Average Point Gain	Math Average Point Gain	Average Family Income
1976	Coached between PSAT and 1st SAT	143	73	70	$29,000
1976	Coached between 1st SAT and 2nd SAT	135	65	70	26,000
1976	Noncoached	60	29	31	21,000
1975	Coached between PSAT and 1st SAT	114	55	59	30,000
1975	Coached between 1st SAT and 2nd SAT	104	47	57	26,000
1975	Noncoached	44	17	27	20,000

SOURCE: National Education Association. *Measurement and Testing: An NEA Perspective.* Washington, DC: NEA, 1980, p. 25. Copyright 1980 by the National Education Association. Used with permission.

vocabulary, and reading skills over a six-week period. The mean gain of the 20 seniors enrolled in the program was 98 SAT-V points. The long-term preparation included a 50-minute daily training from September to March; the mean gain for the 80 students was 109 SAT-V points. According to Slack and Porter, the 20 seniors in the short-term course experienced a gain of 84.7 SAT-V points above the 13.7 point increase normally attributed to practice and growth between tests over a five-month period; the 80 students enrolled in the long-term course experienced a gain of 79.1 points above the 29.9 point increase expected over a 12-month period between tests. [258:158-159]* Additionally, Slack and Porter claimed that ETS failed to reference several studies by Marron in 1965. These studies of high school graduates in a private preparatory school showed that students increased their SAT-V score by an average of 58 points and their SAT-M score by an average of 79 points. These gains were 41.1 points above the gains in verbal scores and 67.6 points above the gains in math scores attributed to practice and growth over the six months between tests. [258:159]

Table 14 shows the results of 29 coaching studies completed prior to 1968. The authors reported that the weighted mean gains of the studies cited by ETS are 16 points for the SAT-V over a control group, compared to 40 points* in studies not cited by ETS. The weighted mean gains of all studies was 29 SAT-V points. Similarly, the studies cited by ETS showed an increase of 12 points for the SAT-M over a control group, compared to 55 points in studies not cited by ETS. The weighted mean gains of all studies was 33 SAT-M points. [258:161]

Slack and Porter stated that the College Board and ETS point out that score increases attributable to coaching are less than the standard error of measurement, yet they do not

---

*See also, "Errata," *Harvard Educational Review*, 50 (August 1980), p. 457.

TABLE 14.--Effects of Coaching on SAT Scores in Studies Cited and Not Cited by Educational Testing Service[a]

SAT-V Studies[b]	Number of Students Coached	Points[c] Gained over Controls	Significance[d] Level	SAT-M Studies[b]	Number of Students Coached	Points[c] Gained over Controls	Significance[d] Level
*Pallone: Short-term	20	85	<.001	*Marron: Group 1	232	75	<.001
Long-term	80[h]	79	<.001	Group 2	405	66	<.001
*Marron: Group 1	83	59	<.001	Group 3	78	56	<.001
French: Identical items	173	47	<.01	Dyer: No senior math	74	29	<.01
*Marron: Group 2	600	40	<.001	Dear	71	26	<.01
Group 3	5	29	n.s.	French: SAT coached B	161[e,f]	18	<.01
Group 4	26	15	<.001	French: Identical items	165	15	<.01
French: SAT coached	161[e]	18	<.01	Lass	38	11	Not reported
*Roberts and Oppenheim	154	14	<.05	Frankel	45	9	n.s.
Whitla	52	11	n.s.	*Roberts and Oppenheim	188	8	n.s.
Frankel	45	8	n.s.	French: SAT coached A	161[e,f]	6	<.01
French: Vocab. coached	110[e]	5	<.05	Dyer: Taking senior math	165	3	<.05
Dyer	239	5	<.05[g]	Whitla	50	- 7	n.s.
*Coffman and Parry	19	4	n.s.				
Lass	38	3	Not reported				
Dear	60	- 2	n.s.				
Weighted Mean Gains				Weighted Mean Gains			
All studies		29 (median = 16)		All studies		33 (median = 15)	
Studies cited by ETS		16		Studies cited by ETS		12	
Studies not cited by ETS		40[h]		Studies not cited by ETS		55	

---

[a] All studies listed in Table 1 were published prior to the 1968 publication, *Effects of Coaching on Scholastic Aptitude Test Scores* (CEEB, 1968). Studies in bold face were either done by ETS or sponsored by them or the College Board. Asterisked studies are those not cited by ETS and the College Board when they reached their conclusion ". . .slowly and with care. . ." that coaching ". . .is at best likely to yield insignificant increases in scores" (CEEB, 1968, p. 8).

[b] Some studies reported in the same publication have been treated independently. Pallone's short-term and long-term studies (1961) used different students at different times of the year. Marron's study (1965) was carried out in 10 separate schools, all with different programs and students. Marron grouped the schools statistically according to differences in initial and final test scores; we have treated each group as an independent study. In Dyer's study (1953), seniors taking math and not taking math performed differently, and we treat them independently. French (1955) used two experimental schools with different programs, "SAT Coached" and "Vocabulary Coached," which are treated independently. Students taking senior math and not taking senior math did not differ in performance in either the French (1955) or Dear (1958) studies and are not treated independently in the table. Roberts and Oppenheim (1966) reported on rural and urban students who did not differ in performance. Coffman and Parry (1967) reported on three groups. We have omitted Group C because the class period "was not long enough to permit the administration of the total verbal test" (p. 293), the incomplete scores were equated to the College Board scale using a different group of students, and a mean decrease of 28.9 points suggested problems with student motivation. Groups A and B were combined because there was no difference in their performance.

[c] For studies without control groups (Pallone, 1961; Marron, 1965; and Coffman & Parry, 1967), the reported gains in scores have been reduced by the increase in scores to be expected from practice and growth. This expected increase, which varies with the score obtained, was calculated individually for each study (see Footnote 5[i]). Thus, the scores shown are estimates of the gains in points that could be expected over a control group with equivalent initial scores on the SAT.

[d] See Footnote 9[i].

[e] In French's original article (1955) the following n's are given: SAT Coached = 188, Vocabulary Coached = 126. We have used the values from French and Dear (1959) because they appeared in a refereed journal after French's original publication.

[f] SAT Coached A is a comparison between the SAT Coached Group and the Control Group; SAT Coached B compares the SAT Coached and Vocabulary Coached Groups.

[g] See Footnote 3[i].

[h] See "Errata," *Harvard Educational Review*, 50 (August 1980), p. 457.

[i] Footnotes refer to notes in Slack and Porter's original article.

SOURCE: Warner V. Slack and Douglas Porter. "The Scholastic Aptitude Test: A Critical Appraisal," *Harvard Educational Review*, 50 (May 1980), p. 161. Copyright 1980 by the President and Fellows of Harvard College. Used with permission.

discourage college admissions committees from making admissions decisions based on test scores that are less than the standard error. [258:162] Moreover, they reported that ETS studies at eight coeducational colleges in 1962, 1964, and 1965-67 conducted by Schrader and Stewart (1971) found that the difference between the scores of all applicants and those admitted to college was 24 points on the SAT-M and 27 points on the SAT-V, much less than the score gains reported in several coaching studies. [258:160] Overall, Slack and Porter said that they found nine conclusions in publications sponsored by ETS or the CEEB and seven assertions that test score gains attributable to coaching were not significant, although there was ample evidence to the contrary. [258:163] They concluded that students can successfully prepare for the SAT, and the more time students devote to special preparation, the greater their score increases will be. [258:164] "Common sense as well as experimental evidence," Slack and Porter said, "belies the contention that students cannot effectively study for the SAT." [258:163] Slack and Porter called upon colleges to drop the SAT requirement from their application procedures as Bowdoin College and the University of Wisconsin have done. [286:A2]

In response to the Slack and Porter article, George Hanford of the College Board stated that the authors took a partisan view rather than a scholarly approach to their critical review, since they had previously assisted Ralph Nader in preparing the report, *The Reign of ETS*. Hanford noted that the two studies prominently featured in the critical review, Pallone (1961) and Marron (1965), "involved special preparation *over the better part of a year*." [emphasis in the original] Consequently, these studies were *not* short-term drills on test questions, but rather "studies of the effects of extended educational programs focused on the SAT." Hanford further noted that both studies were flawed and categorically rejected Slack and Porter's assertion that the College Board and ETS has suppressed studies that showed coaching to be effective. Hanford stated that both studies were described in an ETS research bulletin on short-term instruction for the SAT (January 1978). This bulletin was also distributed later as a College Board research and development report. [122:1-2]

Alice J. Irby, a Washington-based ETS vice-president, further stated that several of the studies cited by Slack and Porter lacked control groups and "were defective from a design standpoint." Irby explained that the 1965 studies were actually equivalent to "a fifth year of prep school." She added, "That's learning, not coaching, but the authors have not made that distinction. To the extent that people take more courses and put in much more study, then they should do better on the test and should probably do better in college." She criticized Slack and Porter for their implication that students who attend a coaching school will automatically experience large gains in their test scores. [100:5]

In an August 1980 article in the *Harvard Educational Review*, Rex Jackson of ETS responded to Slack and Porter's "critical appraisal" of the SAT. Jackson noted with regret that the *Harvard Educational Review* had been employed "to present a seriously distorted and inaccurate brief on one side of an important public issue." Jackson accused Slack and Porter of using selective portions of published information concerning the SAT to fit their case, out-of-context quotations, half-truths and misinformation, and of failing to provide a fair presentation of the College Board's statements which describe the SAT as a measure of developed academic abilities, rather than as "a measure of fixed capacities." [134.1:382-383] Jackson stressed that all of the studies listed by Slack and Porter have been previously cited in ETS reports, in reviews of both coaching and instructional programs of longer duration (Pike, 1978). [134.1:384] But he noted that programs offering substantial instruction over an extended period of time "have generally not been discussed in publications concerned exclusively with short-term coaching." [134.1:384] Not only did Slack and Porter lump together short-term coaching programs and longer-term educational programs, Jackson said, they increased "the effects shown in their table by including the results of a special study using a posttest made up of items identical to practice exercises on which students had been coached (French & Dear, 1959)." [134.1:385] Jackson further stated that Slack and Porter devoted too much space to studies that were from 13 to 27 years old, while ignoring more recent studies by Pike and Evans (1972), Alderman and Powers (1979), and the Federal Trade Commission, Bureau of Consumer Protection (1979), which are essential to a comprehensive review of coaching research. [134.1:385] Jackson characterized Slack and Porter's "critical appraisal" of the SAT as unfair, inaccurate, and irresponsible. He emphasized that "decades of practical experience in the use of SAT scores, has persuaded. . . educators that the test scores are useful for their intended purpose--to supplement the secondary school record and other information about applicants in assessing competence for academic work in college." [134.1:390]

Following Jackson's article in the *Harvard Educational Review*, Slack and Porter responded to specific issues raised by Jackson in the areas of training for the SAT, prediction of academic performance, and aptitude versus achievement testing. They concluded with a finding made in their original article: "we hope that students will remember that [SAT] scores are unrelated to the quality of their minds." [258.1:399]

Additional Research on Special Preparation Sponsored by the College Board

During the past decade the College Board has sponsored several studies to determine the susceptibility of the SAT to special preparation. Discussed below are the results of studies conducted by Pike and Evans (1972), Pike (1979), Alderman and Powers (1979), and Powers and Alderman (1979).

The Pike and Evans report.-- Lewis W. Pike and Franklin R. Evans conducted a study to determine the susceptibility of item formats on the SAT mathematical section to coaching. Pike and Evans's research bulletin, *The Effects of Special Instruction for Three Kinds of Mathematics Aptitude Items* (May 1972), examined quantitative comparisons (QC) as a possible replacement for regular mathematics (RM) and data sufficiency (DS) formats. The QC format requires examinees to determine the relationship of two quantities (i.e., mark "A" if the quantity in column A is larger; "B," if the quantity in column B is larger; "C," if the quantities are equal; and "D," if there is insufficient information to determine the quantitative relationship). The DS format presents examinees with a question and two statements labeled (1) and (2), with the objective being to determine whether the question could be answered by "A," one alone; "B," two alone; "C," by one and two together; "D" by either one alone or two alone; or "E," by neither statement alone nor by one and two together. The RM format requires examinees to determine which of five possible mathematics solutions given is correct. [225:1-2]

Eleventh grade volunteers from 12 high schools participated in the study. Eight suburban schools were located in the vicinity of Princeton, New Jersey, and two urban schools were in Ohio and two in western Pennsylvania. The volunteers were randomly assigned to two instructional groups and one control group. All three groups were given 21 hours of instruction over seven Saturday mornings and about 21 hours of homework from workbooks. Of the 555 volunteers pretested, fewer than 10 percent dropped out of the program. [225.1:12-16] A different instructional program was developed for each of the three mathematics item formats, and pretests and posttests were administered. [225:51]

According to Pike and Evans, results of the statistical analyses found that each of the three mathematics item formats, QC, DS, and RM, was susceptible to special preparation, especially the complex or novel item formats. A further analysis indicated that the volunteers appeared to have learned basic mathematical concepts and a systematic approach to the complex item format. Pike and Evans reported mean gains of almost a full standard deviation for the groups instructed in the complex item formats. This led them to conclude that these gains were "of practical consequence and likely to influence admission decisions." [225:52]

The Pike report.-- Another study conducted by Pike summarized research literature on special preparation and testwiseness. Published as a College Board research and development report in 1979, *Short-Term Instruction, Testwiseness, and the Scholastic Aptitude Test: A Literature Review with Research Recommendations*, stated that questions regarding special instruction for the SAT were relevant because the continuing importance of SAT scores pressured students to seek "instruction for the SAT," which encouraged commercial coaching enterprises as well as some public and private schools to offer such preparation. [223:5]

According to Pike, except for the differences in the relative period of instruction, both short-term instruction (STI) and intermediate-term instruction (ITI) try to improve test scores through special instruction at an accelerated pace compared to the amount of time normally judged necessary for substantial changes to occur. Testwiseness (TW) allows individuals to show their abilities to their best advantage through knowledge about test taking itself. [223:6-7]

Pike reviewed 10 studies pertaining to the mathematical section of the SAT. Seven of the studies (Dyer, French, Dear, Lass, Frankel, Whitla, and Roberts and Oppenheim) indicated that score averages attributable to STI were not sufficient to justify having students invest time in this type of instruction to improve their test scores. [223:23] These studies reported score changes ranging from slight losses to gains of 20 SAT-M points. In contrast, three other studies (Marron, McCarthy, and Pike and Evans) indicated that overall STI increased SAT-M scores significantly. The Pike and Evans study estimated gains of about 33 SAT-M points, the McCarthy study of about 41 SAT-M points, and the Marron study of about 79 SAT-M points. [223:23] Equally important, a comparison of SAT-M and SAT-V findings for seven studies that reported the effects of coaching and short-term instruction on both (Dyer, French, Lass, Dear, Frankel, Whitla, and Roberts and Oppenheim) found that "SAT-V gains were strongly equivalent to those observed for the SAT-M." The high, median, and low STI effects for the SAT-M were 18, 6, and -9, respectively; the effects for the SAT-V were 18, 11, and 0, respectively. [223:27-28] Overall, the studies had substantially different designs in the number of subjects, the use of control groups, and the use of pretest and post-test data. [223:20] Pike reported, however, that there were meaningful STI effects for the SAT-M, but the STI effects for the SAT-V were "unresolved." [233:3] It should be noted that Pike's literature review included complete descriptions of the Pallone study (1961) and the Marron study (1965), which Slack and Porter contended have not been cited by ETS and the CEEB.

At the conclusion of his literature review, Pike recommended that future research regarding STI, ITI, and TW should attempt to "maximize the fairness and validity of the SAT with regard to its short-term and intermediate-term instruction . . .score components"; "to foster realistic understanding and expectations regarding possible outcomes of STI and ITI," rather than discouraging test preparedness per se; and to pursue "a more basic understanding of the processes involved in test taking and contributing to aptitude test scores." [223:34]

Shortly after the release of his literature review, Pike was dismissed by ETS. ETS said that the dismissal resulted from financial cutbacks and was unrelated to Pike's research. Now employed as a senior research associate at the National Institute of Education, Pike stated that he believes there is a link between the College Board's recent statements regarding "formal instruction for the SAT" and the FTC's study of the effects of coaching on standardized admission examinations. [136:9] "I don't think a student can artificially raise his test scores above his ability through test preparation," Pike commented, "but he might bring his scores up to where his ability really is." [237:1, 20]

The Alderman and Powers report.-- Another recent study commissioned by the College Board pertaining to instruction for the SAT was a February 1979 research report by Donald L. Alderman and Donald E. Powers, *The Effects of Special Preparation on SAT Verbal Scores*. Alderman and Powers examined five public and three private secondary schools that had special preparation programs already in place. [13:7] For their study, 239 students were in a treatment group, and 320, in a control group. [13:12-13] Nearly all of the volunteers were high school juniors enrolled in a college preparatory curriculum and most had some experience in taking verbal aptitude tests. [13:8, 21] In most cases an English teacher was responsible for the instruction, commercial review books were also utilized. [13:25]

Seven of the eight schools studied offered special preparation programs of short duration. The number of hours devoted to special preparation ranged from five to 11 hours, with an average of 8.5 hours of preparation. The eighth school offered special preparation of considerably longer duration (45 hours), which increased the average time devoted to special preparation at all of the schools to 13 hours. [13:25][*]

Results of the Alderman and Powers study revealed that special preparation for the SAT-V ranged from -3 to +28 points at the eight schools.[**] [13:20] The average gain attributable to special preparation was eight points, which corresponds to one additional item answered correctly on the SAT-V. [13:20] They observed that students improved their performance in both analogies and antonyms by one-half item each, but the effects for reading comprehension and sentence completion were negligible. [13:18] In light of these findings, Alderman and Powers said that coaching programs directed toward the SAT-V test were "an ineffective response to declines in student scores on the test." [13:21] Nonetheless, the authors stated that it would be inappropriate to make inferences based on their study for student populations such as minority groups, those who lacked PSAT experience, and those attending smaller secondary schools. Rather than adopting special test programs directed toward the SAT, Alderman and Powers recommended that communities and schools concerned with student performance on the SAT should "strengthen the regular curriculum." [13:21]

---

[*] It should be noted that many special preparation courses offered by commercial coaching enterprises typically offer preparation for the SAT some two or three times longer in duration than seven of the school programs reported by Alderman and Powers. Of 22 commercial coaching schools for which data were available, as reported by the Boston Regional Office of the FTC, the average number of hours devoted to special preparation was about 29 hours. Stanley H. Kaplan (School A in BRO's study) offers 40 hours of classroom instruction and Test Preparation Center (School B) offers 24 hours of special preparation. [90:37-40]

[**] According to Slack and Porter, "The large school-to-school variation in points gained must reflect differences in the quality of instruction." [258:160]

The Powers and Alderman report.-- In order to provide secondary schools and students with more information, the College Board has developed a descriptive booklet, *Taking the SAT*, which replaced an older publication, *About the SAT*. The new publication is designed to familiarize candidates with the SAT and the Test of Standard Written English, enhance test taking skills, reduce test anxiety, and help students become more self-confident. The 48-page booklet contains a complete sample SAT, test answers, directions for estimating SAT scaled scores, and strategies for guessing and using time wisely.* [227:1]

In a February 1979 research report titled *The Use, Acceptance, and Impact of Taking the SAT: A Test Familiarization Booklet*, Powers and Alderman reported the results of a survey of student opinions regarding the usefulness of *Taking the SAT* and attempted to estimate the effects that the booklet may have had on test taking and test scores. [227:1] Prepublication copies of *Taking the SAT* plus other materials regularly sent to test registrants were mailed to half of the random sample of 2,024 juniors, while the remaining half received the materials regularly sent to registrants plus *About the SAT*, rather than the new booklet. [227:1-2] Of the students responding to the Powers and Alderman survey, about 38 percent reported reading *Taking the SAT* completely at least once, another third read most of it, and 28 percent either skimmed the publication or read parts of it. Less than three percent made no use of the booklet at all. With regard to the sample test, Powers and Alderman found that it was not used as widely as the text. About 77 percent of the students reported attempting at least part of the test, while 36 percent actually completed it; 23 percent did not undertake any of the questions. [227:3-4] Nearly all of the experimental students (96 percent) felt that *Taking the SAT* had heightened their awareness of what to expect on the SAT. Most students (76 percent) also felt that they had gained more self-confidence because of the booklet. But a comparison of the test scores of those who were sent the new booklet to the test scores of the random sample which did not receive the test booklet showed no differences in the test scores of either group in the verbal or mathematical sections of the SAT or on the Test of Standard Written English. [227:24-25] Given the strong interest among secondary schools in special preparation, Powers and Alderman recommended that "school personnel might wish to consider group administration of the sample test as an alternative to encouraging students to undertake the test on their own." [227:26]

---

*Students who receive *Taking the SAT* are admonished that, "Drilling or last-minute cramming probably will not do much to prepare you for the test." [11:3]

Summary of the Special Preparation Controversy

The initial study by the Boston Regional Office of the Federal Trade Commission, *The Effects of Coaching on Standardized Admission Examinations* (1978), found statistically significant differences between the SAT test scores of coached and uncoached students. [90:1] It was estimated that students attending a commercial preparation course who were coached four hours a week over a 10-week period could improve their total SAT verbal and mathematical scores by as much as 100 points. [188:102] Citing major deficiencies in the BRO study but using the same data base, a report by the Bureau of Consumer Protection of the FTC (1979) found that a certain class of students identified as underachievers, or students who tend to do poorly on standardized tests in light of their previous grades and class rank, could improve their SAT verbal and mathematical scores by an average of 25 test points in each section of the test. [89:1] A reanalysis of the 1979 FTC data by the National Education Association showed that students who received coaching between their PSAT and first SAT or between their first and second SAT experienced score increases between 104 and 143 points, compared to an uncoached group. [194:25] According to Slack and Porter, ETS's studies of special preparation prior to 1968 indicated that coached students could raise their scores by 16 points on the SAT-V, compared to 40 points in studies not cited by ETS. Similarly, in studies of the SAT-M cited by ETS, students could raise their test scores, on the average of 12 test points, compared to 55 points in studies not cited by ETS. [258:161]

On the other hand, ETS's reanalysis of the FTC data found the effects of coaching to be less than both the 1978 BRO and 1979 BCP studies due to the combined effects of coaching and self-selection. After taking into account differential growth and motivation, ETS observed that the effects of special preparation were approximately 11 points for the SAT-V and 30 points for the SAT-M. [177:6] ETS stated that the findings for commercial coaching were consistent with prior studies on the effectiveness of coaching. Moreover, score increases of 20 to 30 points on the College Board's 200-to-800 point scale were equivalent to about three additional test items correct, a gain which students could expect to obtain by simply devoting additional time to their academic courses. [177:7] Other publications and research sponsored by the College Board pointed out that schools and communities interested in helping students prepare for the SAT would benefit by strengthening the existing curriculum and administering the practice test contained in the CEEB publication *Taking the SAT*. [13:21; 227:26]

Despite previous research, or perhaps because of these widely divergent findings, the coaching controversy is not expected to subside quickly. In fact, there is evidence that interest in coaching or special preparation for standardized college admissions examinations is increasing rapidly. According to the Boston Regional Office of the FTC, an industry-wide structure now exists which would allow coaching schools to reach the 2.5 million individuals who take standardized admissions examinations annually, with an estimated sales potential of one-half billion dollars. [90:35-36] Additionally, some colleges and universities are now offering test preparation programs for students. St. John's University in New York offers a $125 coaching course for students planning to take the SAT. [247:2] Rutgers University offers a special preparation program for students planning to take the Law School Admission Test. [237:1, 20] Harcourt Brace, Jovanovich, Inc. a $300 million conglomerate, has acquired a coaching school. [90:36] Citibank has hired the Sexton School to prepare its employees for the Graduate Management Admissions Test and the LSAT. [237:1, 20] NEA has prepared a standardized testing kit for its members, which contains a list of recommendations for parents interested in finding special preparation for their children. [201:1] The National Institute of Education has indicated an interest in planning a conference to assess the effects of special preparation on students who take the SAT. [247:2] The Federal Trade Commission is continuing to monitor the advertising claims of commercial coaching enterprises. [112:2] In 1979 the Department of Health, Education, and Welfare set aside $3 million to the Health Resources Administration Office of Health Resources Opportunity to permit disadvantaged students interested in health professions to be coached at 21 universities. [237:1, 20] According to an ETS survey by Alderman and Powers, nearly one-third or 54 of 175 public and private schools in seven northeastern states have indicated that they offer special courses, electives, or extracurricular activities for students who wish to prepare for the verbal portion of the SAT. [13:4]

THE TRUTH-IN-TESTING CONTROVERSY: CONSUMER RIGHTS VS. GOVERNMENT INTRUSION

Truth-in-testing legislation is essentially an issue of consumer rights versus government intrusion in education. This legislation is a response to consumer desire for more information about the uses and results of standardized college admissions tests. Proponents of truth-in-test legislation assert that they should have the same basic right in obtaining test scores, test questions and answers, and studies related to college admissions as they have in obtaining information contained in school files, credit reports, law enforcement records, and government files. [226:1] These test consumers argue that because of the ambiguity and mystery surrounding standardized testing, they should have the right to examine tests to insure that scores are accurately computed and reported. [152:1] Legislation mandating the disclosure of test questions and answers after an examination is necessary, they argue, in order for people to know how they are being rated and judged. [98:11] Advocates of truth-in-testing believe that legislation will provide the following:

1. Increase the public's understanding of intent and limitations of tests; [130:5]
2. Alleviate bias and promote accountability in the standardized admissions process; [292:2]
3. Permit students to learn from reviewing their mistakes and provide test-taking assistance to those who cannot afford to attend expensive coaching schools; [292:3]
4. Reduce the risk of computer error causing a test-taker to be incorrectly scored; and [292:2]
5. Improve the quality of both the tests and public debates surrounding them. [221:3]

Advocates of testing legislation perceive it as strengthening their "right-to-know" and making the testing agencies "accountable" for a service purchased by millions of individuals annually. [28:2, 226:2]

Test sponsors and test agencies, on the other hand, have expressed considerable alarm over laws which mandate the release of test questions and answers to the public. Measures calling for the disclosure of secure tests, they argue, are an example of unwarranted government intervention into educational affairs and an attack on the entire college admissions process. [273:4] Testing agencies emphasize that truth-in-testing legislation will make it nearly impossible for them to continue their current programs and level of services. [53:1] The disclosure of secure test questions and answers, moreover, will irreparably harm the quality of the tests. As a result, many colleges and universities may lose confidence in the ability of the tests to predict the future academic performance of students. [138:19] Test sponsors and test agencies contend that disclosure will:

1. Destroy the security of each test and make reuse impossible; [46:1]

2. Increase costs of test development and reduce services to clients and students, [46:2];
3. Pressure teachers to teach to the test and increase coaching activity, [46:3];
4. Ignore the diverse needs of test takers who must have access to examinations at different times of the year, [72:2];
5. Lead to limitations of educational opportunities for minority and disadvantaged students, [46:4]; and
6. Violate test organizations' Fourteenth Amendment rights, as well as other provisions of the United States Constitution and the New York State Constitution. [52:1]

"In this war, directed against testing as its declared target," stated ETS president William Turnbull, "the traditional approach to intellectual debate has been shunted aside and replaced by political advocacy." [68:2]

Testing organizations have stated that they support more openness in testing in principle, but prefer to implement changes in standardized testing voluntarily. [65:1-2] Hastily enacted testing legislation, advanced under the fraudulent title of truth-in-testing, they maintain, is itself a "consumer fraud." [231:190] For these reasons, truth-in-testing legislation remains a highly controversial issue.

Origins of New York's Truth-in-Testing Law

Although California enacted the nation's first law requiring test publishers to disclose information about a test (September 1978), New York's truth-in-testing law is more comprehensive in scope and broader in influence. Specifically, the New York law requires test publishers to disclose actual examination questions and answers to test subjects, upon request, 30 days after the release of test scores. [33:7] Leaders in the enactment of the New York legislation were Senator Kenneth P. LaValle (R), Chairman of the New York Senate Committee on Higher Education, and Assemblyman Albert Vann (D). [175:1]

The legislation was the subject of intensive lobbying during the New York Senate Higher Education committee meetings. One educational news service referred to the impressive array of advocates and equally formidable opponents as the "War of the Worlds." Supporters of LaValle's truth-in-testing measure included the New York State P. T. A., National Education Association, New York State United Teachers, New York Public Interest Research Group, National Association for the Advancement of Colored People, National Conference of Black Lawyers, and the Attorney General of the State of New York. On the opposing side were nearly all major test sponsors, administrators, and organizations, including CEEB and ETS. Many prominent college admissions officers and the New York Commissioner of Education, Gordon Ambach, testified against passage of the proposed legislation. [145:35; 211:4; 36:2]

New York governor Hugh Carey approved the bill, 5200-A, which added Article 7-A to the Education Law, in relation to standardized testing, on July 13, 1979. The law became effective on January 1, 1980. It applies to virtually all college and graduate school admissions tests, including the PSAT, SAT, ACT Assessment, GRE, Medical College Admission Test, Dental Admission Testing Program, and the LSAT. [36:1] The legislation does not apply to any state, federal, or local civil service test; any test used solely for placement, credit-by-examination, or other nonadmission purpose; or any test developed and administered by an individual school or institution solely for its own purpose. [264:1]

The College Entrance Examination Board, the Educational Testing Service, and the American College Testing Program indicated that they intend to comply with the New York law; however, many test sponsors have withdrawn their tests from New York. The majority of these tests are related to the health profession, and include the New Medical College Admissions Test, Miller Analogies Test, Optometry College Admissions Test, Pharmacy College Admissions Test, and the Entrance Examination for Schools of Nursing. [34:45-46] Compared with the large numbers of candidates taking the SAT and ACT, most of the test programs withdrawn from the state are given to small numbers of persons, e.g., six of the tests are administered to fewer than 500 candidates a year. [250:1]

Provisions of New York's Truth-in-Testing Law

The New York law demands full disclosure and requires testing agencies to provide tests for public review. Highlights of its provisions in relation to standardized testing follow.

The test agency must file, within 30 days, a copy of all test questions used in calculating the test subject's raw score; the corresponding acceptable answers to those questions; and all rules for converting raw scores into those scores reported to the test subject, together with an explanation of such rules. [264:2]

Within 90 days after filing a standardized test, the test agency must provide the test taker with the opportunity to secure a copy of the test questions, a copy of the test subject's answer sheet, a copy of the correct answer sheet, and a statement of the raw score used to calculate the scores reported to the test subject. The agency may charge a nominal fee for providing this information. [264:2]

Along with the registration form or score report for a test, each test agency must give the following information:

- The purposes for which the test is constructed and is intended to be used.
- Statements designed to provide information for interpreting test results, including explanations of the test score scale, the standard error of measurement of the test, and a list of available correlations between test scores and grades, successful completion of a course of study, and parental income.
- How the test scores will be reported, whether the raw test scores will be converted in any way before being reported to the test subject, and if the testing agency will use the test score either by itself or with any other information about the test subject to predict applicant's future academic performance.
- A description of any promises that the test agency makes to the test taker regarding accuracy of scoring, timely forwarding of information, and policies for notification of inaccuracies in scoring or score reporting.
- Whether test scores are the property of the test subject; the time period during which the results are retained by the test agency; and clarification of policies regarding storage, disposal, and future use of test score data.

An individual's score shall not be released or disclosed by the test agency to any person, organization, corporation, association, college, university, or governmental agency unless specifically authorized by the test subject. A test agency may, however, release all scores received by a test subject on a test to anyone designated by the test subject to receive the current score. [264:3-4]

Major Issues Vigorously Debated

In an article in *Compact*, Rexford Brown, director of publications and user products for the National Assessment of Educational Progress, observed that truth-in-testing "is a complex issue, touching explicitly or implicitly on many education subissues that have polarized Americans for generations." [33:7] Advocates of truth-in-testing who rally around the principle of fairness, and testing agencies that caution about the unanticipated consequences of hastily enacted legislation, are often arguing at each other rather than with each other, Brown stated. The antagonists seldom draw upon the same facts, examine the same aspects of education, or subscribe to the same beliefs about the function of education. According to Brown, the debate about the need for and the impact of truth-in-testing legislation centers on four major areas: (1) the role and power of testing organizations; (2) the nature and quality of standardized tests; (3) the need for state and federal legislation; and (4) the consequences of full disclosure provisions. [33:7-8] Specific issues that relate to the SAT, ACT, and truth-in-testing legislation follow.

Disclosure of secure test questions and answers.--Critics of standardized college admissions tests maintain that disclosure of secure test questions and answers will increase knowledge about the tests and will permit test subjects to view the criteria used to evaluate their abilities. [266:13] In testimony before the New York State Senate and Assembly Higher Education committees, Lewis Pike, former ETS researcher and now senior associate at the National Institute of Education, stated that college admissions tests are a critical public trust and the assumption that students who want to see their tests must demonstrate "the need to know" was inappropriate. [224:2] Pike strongly endorsed testing legislation, saying that it would "provide a source and motivation for quality control that is simply not otherwise present, however well-intentioned the test producer may be." [222:2]

Among the most ardent supporters of truth-in-testing legislation, the New York Public Interest Research Group (NYPIRG) asserted that the disclosure of test questions and answers will improve the quality of question writing. Subjecting test questions to scrutiny by subject matter experts, educators, and the public will reduce the extensive criticism of the scientific accuracy of test questions, NYPIRG believes. Public disclosure is essential because of the controversy in academic circles surrounding the "correct answers" on standardized tests. NYPIRG cited the release of a secure ETS test in 1973, a Multistate Bar Examination (MBE). Following the disclosure of the test, law professors at Georgetown University and other Washington, D.C., institutions expressed disagreement on the "correct answers" to 25 percent of the questions. According to NYPIRG, ETS and its clients agreed that five of the questions appeared to have more than one correct answer. Sincer 1973 no other tests have been divulged, NYPIRG said. [209:3]

Proponents of standardized testing, however, have claimed that the public is being deceived by false statements which misrepresent the role of test agencies in American education. The public is being misled by loaded words like "truth," "fairness," and "public interest." [118:12] Secure testing programs are not secret, according to ETS, but are essential to insuring fairness. ETS reported that by maintaining the security of tests it can do the following: (1) ensure that all students receive equal treatment and that no one has an unfair advantage by seeing a test in advance; (2) reuse tests and keep down the cost of development; and (3) equate scores between different test editions, thereby assuring that students take tests of similar difficulty. [65:1]

While many testing agencies support the principle of openness in testing, they claim that truth-in-testing legislation is unnecessary since information about tests is widely disseminated and available to both students and institutions. ETS pointed out that it provides sample tests for the SAT, GRE, GMAT, and LSAT, sends scores and an interpretation of their meaning to students, forwards test results to institutions only with the student's permission, summarizes statistics for the press and state policy makers, and publishers research in professional journals and magazines. Additionally, students are asked for advice on how to improve programs and services. [65:1] With regard to NYPIRG's criticism of the Multistate Bar Examination, ETS said that the complaints stemmed from a single law professor writing one article. Since 1972 more than 59 actual MBE questions have been released and are available in book form. Sixty questions from previous tests are in the candidate bulletin. [70:3]

In a statement opposing New York's joint Senate-Assembly Bill 5200, the College Board emphasized that the disclosure of secure tests was probably one of the bill's most harmful features. The College Board said that security measures are necessary in order to prevent students from gaining access to the tests. If test security were broken, then the tests would lose their value and fairness and new tests would have to be created. [53:1] Advocates of testing legislation, the College Board said, failed to understand the nature of the SAT. The SAT measures developed reasoning abilities in English and mathematics; it is not based upon a specific curriculum or a set of courses. Consequently, test takers cannot "learn" by viewing mistakes on their tests. The systematic disclosure of secure test questions, moreover, could harm the quality of the tests because of the time involved in pretesting questions to weed out any possible cultural ambiguity or bias. [46:3-4]

An October 1979 editorial in the *Washington Post* suggested that the underlying premise of testing legislation was that the SAT and other standardized tests victimize poor and minority students. The editorial stated that such tests may not measure the academic potential of some students, poor, minority, and white, because their intellectual growth had been stifled by inadequate elementary and secondary schools. "The fault is not in the tests but in the schools," the editorial concluded, "and good admissions officers take that into consideration." [279:A12]

Commenting on the impact of the disclosure provision, George Hanford, president of the College Board, said there were serious differences of opinion regarding the value of publicizing SAT questions. Hanford added that the release of each form of the SAT after its administration was "wasteful, educationally questionable and of little benefit to students." [51:3] Other representatives of testing organizations and college admissions officers expressed similar sentiments. Fred Hargadon, chairman of the College Board, expressed doubt about the SAT's ability to survive repeated test disclosure without raising "serious questions" about its validity. [138:1] Repeated disclosure of secure test questions, according to Richard Ferguson, vice-president for research and development at the American College Testing Program, may weaken the ACT battery since it has stronger ties to the curriculum. [140:10] According to William Fitzsimmons, director of undergraduate admissions at Harvard, test disclosure might enable some students to boost their SAT scores. Consequently, Harvard planned to look "hard" at the predictive value of the SAT to see if it has declined. Another admissions officer, Lawrence Momo, associate director of admissions at Columbia University, noted that test disclosure would prompt Columbia to examine SAT scores in a different way, but how differently remained to be seen. [138:19]

The full impact of the test disclosure provision of New York's truth-in-testing law is not expected to create concern on college campuses until after fall 1980 recruitment. At that time colleges and universities will have to assess whether students are affected by the SAT's disclosure. Some have expressed concern that unless fears concerning the potential effects of test disclosure are allayed, many colleges and universities may lose confidence in the SAT, regardless of whether its validity is actually impaired. [138:19]

Cost of test development.-- The cost issue was one of the primary objections to truth-in-testing legislation raised by the test manufacturers. But according to Senator LaValle, testing agencies' claims for increasing costs were based on the false assumption that the bill would require the creation of separate tests for New Yorkers. LaValle contended that there was no justification for test agencies to increase the costs of taking tests in New York for tests that are administered nationwide. [153:60] Vito Perrone, dean of the Center for Teaching and Learning at the University of North Dakota, thought that there was a degree of intimidation present in the threats of the test makers to treble costs and reduce test administrations. Perrone said that dire consequences need not prevail unless "the testing industry refuses to bring already existing knowledge and fresh commitment to bear on the new circumstances" of test disclosure. [221:2] Shirley Chisholm (D-N.Y.), co-sponsor of a federal truth-in-testing bill, asserted that "people's lives are more important than cost." [230:A27]

Citing ETS's own internal studies, Allan Nairn and Ronald Brownstein noted that approximately five percent of each student's test fee is applied to test development and between 22 percent and 27 percent of the fee is applied to ETS's profit margin. Due to the New York legislation, ETS has estimated that the additional cost for developing new SAT questions is $1,092,000. According to Nairn and Brownstein, this estimate provides ETS and the College Board with a $2 million profit. [35:189]

NYPIRG also alleged that ETS's estimate of test development costs is highly inflated. NYPIRG claimed that a January 1972 ETS "activity analysis" showed that the actual cost of developing the SAT and Achievement tests was $.32 per test. This amounted to 6.6 percent of the total cost to ETS of administering the test $4.82 and only 5.7 percent of the fee paid by the candidate ($5.75). In contrast, the profit margin taken by ETS and the College Board on these tests was $.93, or 16 percent of the candidate's fee, nearly three times larger than test development spending. [209:4] Table 15 shows how much of a candidate's fee goes to test development, according to NYPIRG.

According to NYPIRG, figures from 1975 and 1976 showed College Board profit margins between 22 and 27 percent. [209:5] Table 16 gives the admissions testing program revenues and expenses from 1974-75 to 1976-77.

Table 17 shows ETS test development spending as cited by NYPIRG. In 1976 ETS spent $291,000 on SAT Achievement test construction expenses. This amounted to 2.5 percent of the 1976 SAT Achievement program income from student fees. In 1977 ETS budgeted $389,000 on test construction, or 3.1 percent of expected program income. Since most test questions are newly written, and have been for almost 30 years, NYPIRG said that there was no apparent lack of high-quality questions available to test makers. Consequently, ETS's claim that test construction costs would escalate because new test questions would have to be constructed seemed highly implausible. [209:3-4]

TABLE 15.--Comparative Analysis of Program Costs by Activity, 1970-71

Test Program	Fee Paid by Candidate	Total Cost (per session Candidate) Incurred by ETS	Cost of Test Development	Test Dev. as % of Total ETS Cost	Test Div. as % of Fee Paid by Cand.
SAT and Achievement Tests	$ 5.75	$ 4.82	$ .32	6.6%	5.7%
LSAT	13.50	10.83	.49	4.5	3.6
GRE	8.00	7.17	.62	8.6	7.8
ATGSB (now GMAT)	10.00	9.22	.44	4.8	4.4

SOURCE: "Activity Analysis," Internal ETS Study, Jan. 31, 1972.

Reported by New York Public Interest Research Group. "Statement of Factual Correction and Rebuttal to Educational Testing Service Memorandum in Opposition." Albany, New York: NYPIRG, May 1979, p. 5.

TABLE 16.--Admissions Testing Program Revenues and Expenses

	Revenues	Expenses	Surplus	Percent Profit
1974-75	$16,036,276	$12,550,541	$3,485,735	27
1975-76	16,260,652	13,232,474	3,028,178	22
*1976-77	17,640,000	14,099,000	3,541,000	25

*projected

SOURCE: College Board Statement of Revenues, Expense and Fund Balance, Year Ended June 30, 1976, with Comparative Figures for 1975 and 1977 Current Projections.

Reported by New York Public Interest Research Group. "Statement of Factual Correction and Rebuttal to Educational Testing Service Memorandum in Opposition." Albany, New York: NYPIRG, May 1979, p. 5.

In response to allegations that test costs are highly inflated, the College Board stated that its historical policy has been to make the SAT widely available to high school students, and at the least possible cost. Testing legislation suggests that students should pay more and forego the greater availability of testing in order to have the opportunity to view their examinations, the College Board said. Additionally, it was noted that there has not been any substantial demand from test candidates, admissions counselors, or officers that would warrant this "trade-off." [46:2-3] Alice Irby, an ETS vice-president asked, "Where is the equity for making all pay for what few may want?" [133:2] In response to the charge of excessive profit, ETS said that the College Board shows a 0.8 percent "profit" (excess of revenue over expenses) for the SAT and achievement tests. This amounted to $155,000 in fiscal 1978. The "22 percent

TABLE 17.--ETS Test Development Spending Breakdown and Budget

		College Board Admissions Testing Program (SAT & Achievement Tests)					
	Test Construction Expenses	Program Admin. Expenses	Test Const. as % of Program Admin. Exps.	Program Admin. Income	Test Const. as % of Program Admin. Income	ATP Candidate Volume*	Test Const. Expenses per Candidate**
1976 Actual:	$291,000	$10,079,000	2.9%	$11,490,000	2.5%	1,737,000	$ .17
1977 Budget:	389,000	10,864,000	3.6	12,396,000	3.1	1,430,000	.27

*Candidate Volume figures are from *ETS in Fact*, an ETS publication.

**Expenses per Candidate calculated by dividing Test Construction Expenses by Candidate Volume.

SOURCE: "ETS Confidential Project Operating Statement by Pos.--Category/Project, Period Ending 3/31/77," a 66-page computer printout.

Reported by New York Public Interest Group. "Statement of Factual Correction and Rebuttal to Educational Testing Service Memorandum in Opposition." Albany, New York: NYPIRG, May 1979, p. 6.

profit" cited by NYPIRG ignores $11.1 million in College Board expense for support services, administrative and other costs, and research and development. [70:1] ETS emphasized that the costs for developing all new tests were substantial and that neither it nor the program sponsors could absorb consequent increases. [71:4] In 1979 ETS reported an income of $94.1 million vs. expenses of $93 million. The surplus of $1.1 million was used to make improvements to its facilities and to purchase new equipment. [70:1]

In a complaint filed with the New York Supreme Court, the College Board stated that in 1978-79 the cost of test development exceeded $70,000 per form. Because old test questions might be excluded from new tests, it was estimated that the cost of test development would increase to more than $83,000. [59:15] The College Board estimated that the annual replacement cost of the 21 forms of the SAT currently used in New York would run approximately $900,000. [53:1-2]

An initial "Impact Report of the Admission Testing Legislation" prepared by the New York State Education Department in October 1979 cited examples of test programs with low candidate volumes. Since the disclosure of secure tests would necessitate the creation of a new test for each administration, the costs of test development could not be amortized over a group of candidates or over a number of test administrations. Consequently, the legislation would make it difficult for smaller testing companies to survive. [132:4-5] In June 1980, however, Governor Carey approved Senate Bill 9142-A which exempts low-volume tests from the disclosure requirements which are administered to fewer than 2,000 test subjects annually, provided that one test is disclosed at least once every three years. [37:2]

According to Rexford Brown and Merle Steven McClung, the actual cost of test development is an insufficient objection to truth-in-testing legislation, if it is "viewed from an ethical perspective." They predict that the cost argument will subside soon and be considered a "solvable secondary problem." [34:29]

Reduced services for handicapped individuals and religious minorities --

Because of costs associated with developing new test forms, testing organizations say that they are being forced to curtail special services and test administrations for handicapped individuals and religious minorities. The College Board has announced that New York students taking the SAT would be tested only four times between January and June rather than the eight regularly scheduled administrations. [51:1] Similarly, the GRE was reduced from six to four administrations, and the LSAT from four to three. [34:46] Before the implementation of the New York law, the College Board reported that handicapped students could be tested on almost any day and at their convenience; now they will have to take the tests on regularly scheduled dates. Hanford said he was distressed about the cutback in services, but the new law made flexible dates and special arrangements for handicapped test takers almost impossible. [41:1]

According to the New York State Education Department's "Initial Impact Report of the Admission Testing Legislation," the American College Testing Program intended to eliminate residual (on-campus) testing which would affect many adults who need test scores for placement rather than for admission. In 1978 nearly 2,200 students were tested in this manner. The number of special non-Saturday test dates also would be reduced, affecting some 300 students. [132:8]

It has been estimated that the cancellation of Sunday administrations of the SAT would affect approximately 1,800 Sabbath observers in 1980. [132:11] Special make-up administrations of the SAT would probably be eliminated, affecting about 350 to 750 students annually. Because of the high costs of converting tests into Braille, large type, and audio-cassettes, services for approximately 400 students would be restricted to a limited number of test dates. [132:12] Since there are only a few editions of the SAT in Spanish (Prueba de Aptitud Academica), the College Board did not feel it could disclose even one form. As a result, this test has been withdrawn from New York, affecting some 400 students annually. [132:11] The New York State Education Department's report warned that the apparently minor changes in test availability could have a serious impact on the large population of potential candidates. [132:12] The report emphasized that the reduction in special services by test sponsors might result in delaying a candidate's enrollment for a semester or a year, denying credit or advanced standing to undergraduates seeking admission and placement, overburdening some test administration centers on a particular test date, and reducing special services for late registrants and "walk-in" candidates. [132:14]

Yet advocates of truth-in-testing, such as Governor Carey maintain that testing legislation need not result in an "automatic reduction" in the availability of test services to New Yorkers. [36:2] Steven Leifman, national director of the Coalition of Independent College and University Students, claimed that by reducing special services and putting pressure on New York, the College Board may be "trying to substantiate their fears and problems they see with a national law." [41:1]

In February 1980, the New York State Assembly approved a bill requiring the College Board and other test sponsors to give Sunday examinations for religious minorities who observe the Sabbath on Saturday. There was "no need" for test sponsors to take punitive actions against Orthodox Jews and Seventh Day Adventists, said Assemblyman Sheldon Silver, (D-Manhattan) sponsor of the measure. [205:5] In June 1980, Governor Carey signed into law Senate Bill 9142-A which exempts from test disclosure requirements test forms used for make-up purposes and for walk-in, handicapped, and Sabbath observer test takers. Another bill signed by Carey requires that special administrations of standardized tests be provided with the same frequency as regular administrations for test takers who are unable to attend on previously scheduled dates because of religious observance. [37:2]

Errors in correcting and reporting test scores.--

Critics of standardized college admissions tests cite the accuracy in correcting and reporting test scores as an additional reason for the passage of state and federal truth-in-testing legislation. Despite the control measures of testing companies, clerical and computer errors do occur, they maintain. Senator LaValle pointed out that in 1978 over 360,000 New Yorkers were administered standardized admissions examinations, yet these students had no way of knowing whether their tests were accurately marked or of knowing what their mistakes were. [151:2] Proponents of testing legislation believe that test disclosure laws will permit students to verify their scores and further reduce the risk of computer error. [292:2]

In testimony before LaValle's Higher Education committee, Julia Stromsted, legislative director of the Independent Student Coalition, reported that errors in scoring and reporting test scores may adversely affect thousands of test takers annually. Stromsted cited ETS's identification of "unacknowledged repeaters" on the LSAT, in which, several years ago, students were asked on the LSAT application form if they had taken the test before. This question was used to give students the opportunity to acknowledge that they were repeating the test. A check by ETS, using its records on prior candidates, however, identified repeaters who did not answer the question or who gave false answers. The test scores of these students sent to colleges and universities carried the designation of "unacknowledged repeater." According to Stromsted, a computer error in 1975-76 accidentally identified thousands of test candidates as "unacknowledged repeaters." ETS told law schools to disregard the designation, but did not correct the error or inform the test applicants, Stromsted claimed. Another computer error, which affected more than 50,000 candidates, occurred during the 1977 administration of the Graduate Management Admission Test. Incorrect scores were sent to hundreds of graduate schools nationwide. While the error was small, ETS admitted that it could have affected students applying to schools with cutoff scores, Stromsted stated. [153:62]

Prior to the enactment of testing legislation, it was estimated that less than one-tenth of one percent of the 1.5 million SAT takers nationwide make a request to have their scores verified. Of the 1,000 to 1,600 test scores that are reviewed, about 40 are changed, according to Jerry Murphy, director of corporate quality assurance and data policy administration at ETS. According to one education writer, this represents an "error rate" of .0027 of one percent. [293:20]

In testimony before the Maryland Constitutional and Administrative Law Committee, Alice Irby of ETS, testified that verification of scores is possible through present procedures. Answer sheets that are pulled for extensive quality control checks by hand rather than by machine indicate that the accuracy rate is very high. However, if any persons have reason to believe that an error has been made on their test score, they may request rescoring by hand. Irby noted that the Law School Council has had this service for several years, but few students use it. [133:4]

To further insure that test answer sheets are correctly scored, the College Board announced in April 1980 that it would allow the 1.5 million juniors and seniors who took the SAT to buy back their answer sheets and the scoring key, but not the test questions, for a nominal fee of about

$4.00. In New York State, however, students can also buy back the test questions, as well as the answers, for $4.65. [268:A5] The College Board announced that it took this action "to meet consumer concerns over testing without compromising test quality or educational values." [96:8] Furthermore, the College Board said it plans to publish an actual SAT used during the previous academic year for students and researchers to examine, along with information about scoring, standard error of measurement, and each question's difficulty. [268:A5]

Similarly, after the 1980 fall administration of the PSAT nationwide, one million juniors will get back the test questions and answers at no additional charge. Currently the PSAT costs $3.25 and is a retired form of the SAT. [268:A5] Hanford said that these measures were intended to carry out part of the public interest principles that test sponsors and test organizations agreed to last December. [96:8] He said the focus of the new information would be on the PSAT "because it is an early guidance tool." [96:8]

As of September 1980, it had been estimated that 4.8 percent of the 118,000 SAT takers in New York had requested copies of the questions and correct answers along with their own answers; this compares to about 17 percent of those taking the LSAT. [38:250; 94.1:1]

Pro-testing legislation groups believe that the College Board's decision to return test questions and answers to test takers is merely an attempt to halt the enactment of further test disclosure legislation. According to the National Education Association, giving students a way to get back their SAT answer sheets is a small concession. NEA President Willard McGuire said it was "only the 'tip of the iceberg' in problems intrinsic to the testing process." McGuire urged the adoption of truth-in-testing legislation at the state and federal level. [270:2]

Equating the tests.-- In the initial stages of the debate between pro- and anti-testing forces, the issue of equating the tests was highly controversial. [153:60] After much debate, the New York legislature specifically exempted questions used in equating tests from the test disclosure provisions. Currently, only those questions used to calculate the student's raw score, which determines his or her grade, must be disclosed. [209:1]

ETS stated that the equating process is a highly specialized field not even fully understood by many psychometricians. [67:1] The method of equating the SAT involves inserting a section of questions that have been used previously in one or more new forms of the test, with each SAT composed of approximately 20 percent equating questions. [176:2] This process enables test makers to "tie" different forms of the SAT to a standard score scale and it is designed to insure that a student's score of 550 is equivalent to another student's score of 550, throughout the nation and over time. [53:1; 176:2] Although New York State exempted test questions used to equate tests from the disclosure provision, ETS indicated that other state legislatures might not make this exemption, thereby jeopardizing the quality of the tests. [133:3]

Disclosure of background and statistical data.--

New York's truth-in-testing law currently requires test companies to make public after January 1, 1980, all background reports and statistical data pertaining to tests. [264:2] Advocates of truth-in-testing argue that this provision is necessary if scholars, consumers, and educators are to adequately assess test characteristics. NYPIRG has claimed that ETS is withholding several reports. The reports under dispute are "Test Analyses," which is needed to verify the validity of ETS's equating procedures; a major internal study, "Cultural Bias in Testing: Challenge and Response"; and a study on the use of the LSAT by law schools. [209:6]

ETS, on the other hand, claims that the disclosure of background reports, statistical data, and validity studies developed by and for individual colleges and institutions is overly broad in scope. Studies performed under contract for client organizations such as the Law School Admissions Council and the federal government are the property of those entities, and these studies may contain confidential information which ETS does not have the authority to release. Furthermore, much of ETS's statistical data are associated with its tests and designed for internal use. [176:3] On a practical level, the College Board reported that the release of all background reports and statistical data would inundate the Commissioner of Education's office with file cabinets of SAT data alone, not to mention computer printouts of test item analyses; mathematical equations used in scaling and equating tests; and reports concerning the operation of hundreds of test centers, all at an additional cost to the test sponsor. Strict compliance would be "extraordinarily burdensome" and of questionable benefit to anyone, according to the College Board. [53:2] The disclosure of background and statistical data would affect primarily the validity studies provided to institutions which help colleges refine their admission procedures. Rather than being placed in a possible legal confrontation between the law and the institution's rights, many test sponsors indicated that they would stop providing validity studies. The loss of validity data, in turn, would hinder the selection process and affect the applicants, concluded a report by the New York State Education Department. [132:15]

Extraterritorial application.-- Several different provisions of the New York standardized testing law have evoked questions of interpretation. One of those questions involves the issue of extraterritorial application of New York's law. According to Arnold Bloom, New York state education department public information director, the department considers all students applying to institutions in the state to be covered by the law. [233:5] As the law was originally written, any out-of-the-state student applying to an institution in New York could receive his or her test. [294:137] Strict adherence to this interpretation would mean that the College Board would eventually have to reveal all of the test forms it currently uses nationwide.

In response to this interpretation, the College Board filed a complaint in November 1979, asking the New York Supreme Court to declare the law inapplicable to extraterritorial tests. [52:1] The College Board asserted that extraterritorial application of the law violates its due process rights under the Fourteenth Amendment and that it conflicts with both laws in other states and with laws which protect the College Board's property interests in the tests. [59:11-12; 52:1]

Application of the law to tests administered outside New York, the suit stated, would seriously affect the College Board's testing programs and services. Specifically, this interpretation would require them to do the following: (1) disclose virtually all SAT questions, including questions used solely in tests administered outside of New York; (2) discontinue the practice of providing the score reports of seniors in Connecticut, Indiana, and Pennsylvania to their state's scholarship program, unless the student indicated otherwise; and (3) file confidential reports that the CEEB prepares for high schools, colleges, universities, and scholarship sponsors located outside New York. [59:12-13] Hanford commented that "nothing in the language of the law indicates that it was intended to apply to tests given outside of New York." [52:1]

In June 1980, Governor Carey approved the senate bill which amended the education law in relation to standardized testing. He said the bill would clarify truth-in-testing provisions and would limit the law's application to tests administered in New York State. [37:1-2]

Nationwide Movement Predicted

Since the enactment of New York's truth-in-testing law in January 1980, nearly 20 states have introduced bills to regulate standardized admissions tests for college or professional school. This widespread legislative attention is evidence of an emerging trend. [236:189] Steve Solomon, coordinator for NYPIRG's campaign for testing legislation said, "the fact that so many states are considering legislation indicates many students and parents are concerned about the lack of information." [181:1] Currently, only New York and California have enacted laws successfully, but it is estimated that nine other bills have a chance of survival. [29:1, 25] Highlights of proposed state legislation mandating the disclosure of tests and test related information follow.

California: S.B. 2005, signed into law September 1978 and effective in 1979, was the nation's first law requiring test publishers to disclose information to test takers. This law applies to any standardized test used for postsecondary admissions tests given to more than 3,000 students annually. Tests designed for placement, guidance counseling, research, and meeting secondary school graduation requirements are exempt. Test agencies are required to file with the Commissioner of Postsecondary Education a representative sample of test questions and answers equivalent to those used on secure tests. Test sponsors must also disclose the total amount of fees received and the expenses attributable to the test. S.B. 2005 is a limited disclosure law since test takers are provided with information about the tests and their uses but not with test questions and answers. [34:18; 49:2] The California legislature is currently considering an amendment which would provide test takers with a statutory cause of action against any testing agency for unnecessary delays in reporting scores or for the loss of scores. [236:188]

Colorado: S.B. 320, like the California law, applies to standardized tests used for postsecondary admissions administered to more than 3,000 students annually. It is not a full disclosure bill either, and only background and explanatory material must be disclosed to test takers. Tests used for placement, guidance counseling, research, and meeting secondary school graduation requirements are exempt from the legislation. Introduced in January 1979, the Senate Committee on Education postponed it indefinitely. [34:19; 49:2]

Florida: H.B. 1169 applies to any standardized test used for postsecondary admissions, as well as for financial aid for placement. This bill requires full disclosure of test questions and answers and other test information to test takers. Civil service and job placement tests are excluded. Introduced in early 1979, the bill's author withdrew it to amend its original provisions. [34:20; 49:2]

Indiana: S.B. 102 is aimed at postsecondary students. Testing companies are required to report test answers, scoring procedures, potential use of test scores, and other information to both postsecondary education institutions and to students. Violators may be fined up to $500. [257:2; 281:194]

Maryland: H.B. 1425 is a limited disclosure bill that applies to any standardized test used for postsecondary admissions or placement. Test takers were to receive explanatory material, information regarding cutoff scores and material contrasting score use and weight with grades.

Although this bill did not pass, the Maryland House of Delegates is considering other measures which have been introduced recently. [34:21; 49:2; 133:1]

Massachusetts: S. 237 applies to all commercially available standardized achievement or aptitude tests administered in grades K-12. Within 60 days after a school receives test results, parents, students, or both will be permitted to review the corrected responses to tests, but they are not allowed to remove the test items or answers from the school premises. S. 238 focuses on postsecondary and professional school admissions. It requires test companies to submit to the chancellor information on questions used in calculating raw scores, a copy of the answers, and the scoring rules. S. 264 calls for a special commission to undertake an investigative study of the CEEB, its administration, the accuracy of its testing program, and its role in the admissions process. [257:1; 281:194] Massachusetts also requires the disclosure of its competency tests. [33:7]

New Jersey: S.B. 3461 applies to any test developed for the purpose of selection, placement, classification, graduation, or any other reason concerning elementary, secondary, postsecondary, or professional school. This bill calls for the release of test questions, answers, and scoring rules within 30 days after the test results are released. The measure prohibits a test company from charging for the disclosure service, and test scores cannot be forwarded to institutions without the test takers' consent. [34:24; 257:2]

Ohio: H.B. 636 is modeled after the New York law. It is a full disclosure measure that applies to any standardized test for postsecondary admissions selection. Test agencies must fulfill requests for test questions and answers, answer sheet, and raw score. At the time of registration, the agencies must provide the test taker with information about the purposes of the test, subject matter, knowledge and skills being measured, explanation of the score scale and standard error, correlations between test scores and grade, how the scores will be reported, and who owns the scores. Introduced in May 1979, the measure has been assigned to a subcommittee for further debate. [34:22; 49:2]

Pennsylvania: S.B. 994 is also modeled after the New York law. It was referred to the Senate Education Committee on October 2, 1979. Legislation applies to any examination used for admission to undergraduate, graduate, or professional school. All questions used to calculate the raw score and the answer sheet are to be released to test takers upon request, but test items essential for field trials and comparability are exempt. Test agencies must also divulge reports and data prepared for individual institutions. [34:23; 49:2]

South Carolina: A recently introduced measure would require test companies to file postsecondary test questions, answers and other information with the state education department. [281:194; 257:2]

Tennessee: H.B. 1460, the Standardized Testing Act of 1980, covers postsecondary admission tests. It allows test takers to find out what questions they answered correctly. The College Board Achievement Tests and the GRE Advanced Tests are exempt. [257:2]

Texas: H.B. 59 applies to any examination used to select candidates for admission to undergraduate, graduate, or professional school. This bill requires test agencies to provide all test takers with test questions and answers, whether or not they request them. Also, upon written request, all information contained in the testing agencies' records about the test taker must be disclosed. This full disclosure measure was postponed in 1979 by the House Committee on Higher Education. [34:23; 49:2]

Similar legislation has been introduced, postponed, or defeated in Rhode Island, Michigan, Louisiana, Connecticut, Illinois, Minnesota, Mississippi, Missouri, and Oklahoma. [63:7; 29:1, 25]

In their publication *Searching for the Truth about "Truth in Testing" Legislation*, Rexford Brown and Merle Steven McClung observed that most testing measures apply to postsecondary admissions testing and require full or partial disclosure of test questions and answers shortly after the tests are administered. [34:ix] Test legislation laws typically exclude civil service examinations, tests for research and guidance counseling, secondary school graduation requirements, College Board Achievement Tests, and GRE Advanced Tests. [34:17-24] Brown and McClung reported that most state proposals entrust enforcement of the laws to the state's postsecondary education agencies, such as the Pennsylvania Department of Education, Ohio Board of Regents, California Postsecondary Education Commission, and Maryland Board for Higher Education. The different characters and responsibilities of these agencies, Brown and McClung maintained, "suggest a potential for state to state miscommunication and confusion if more laws are enacted." [34:15] Similarly, Dario Robertson, writing in the *Journal of Law and Education* (April 1980) noted that if the current state legislative trend continues, disparate truth-in-testing laws may seriously burden the testing agencies' attempts at compliance. [236:193]

ETS has urged states to postpone the enactment of laws that regulate admissions testing in order to assess the New York experience, particularly the economic impact and the estimated improvement due to unlimited test disclosure; to await legal rulings since some of New York's test disclosure provisions are being challenged on constitutional grounds; and to consider independent studies by the American Council on Education,* the Ford Foundation, and the National Academy of Sciences. [71:2]

In March 1980, in testimony before the Maryland House of Delegates Constitutional and Administrative Law Committee, Alice Irby stressed that the regulation of admissions testing was premature. She stated that testing legislation is "unnecessary and that the objectives of those who want test agencies to provide more information to students about their scores and about the tests themselves are being accomplished through negotiation, re-examination, and deliberation without imposing the long arm of the state." [133:5]

Federal Truth-in-Testing Legislation

Congressional legislation over the merits of truth-in-testing has been under consideration since 1977. The first federal truth-in-testing bill, although unsuccessfully introduced in the House on April 29, 1977, formed the basis for current proposals. [236:190] In 1979 two additional federal bills were introduced, the Gibbons Bill (H.R. 3564) and the Weiss Bill (H.R. 4949). [33:7] The Gibbons Bill, introduced on April 10, 1979, applied to all standardized admissions testing, including aptitude and achievement tests, state licensing examinations for doctors, lawyers, plumbers, and cosmetologists, and civil service examinations. [34:16; 275:2] Sam Gibbons (R-N.Y.) said that his measure was needed to "strengthen consumer rights." Provisions of the bill called for the release of information about test content and scoring, and the prohibition of post-examination test score curving. [98:11] The bill also required test publishers to reveal the required "cut off" scores for higher education institutions. It was not, however, a full disclosure measure since it did not provide for the release of test questions and answers to the public, nor did it require test companies to file anything with a federal agency. This proposal received little attention. [34:16]

Commenting on the Gibbons Bill, Fred Hargadon, chairman of the College Board, stated that it had

---

*The American Council on Education helped found ETS in 1947. The president of ACE serves *ex officio* on the ETS Board of Trustees.

no control over the people who use test scores and that it did not even recommend cutoff scores. He said that the Gibbons measure defined the idea of "test" too broadly and that there was no reason for government regulation of occupational testing. [275:2]

A second federal truth-in-testing bill was introduced by Representative Ted Weiss (D-N.Y.) on July 24, 1979. Known as the "Educational Testing Act of 1979," the measure was cosponsored by Representatives Shirley Chisholm (D-N.Y.) and George Miller (D-Cal.). Referring to the bill as an extension of legislation enacted in California and New York, Congressman Weiss said that . . . the bill "is designed to alleviate bias inherent in the tests, to improve public accountability and to mandate financial disclosure." [292:2]

Like its New York counterpart, H.R. 4949 applies to any standardized test used in the postsecondary admissions process and to organizations that develop, sponsor or administer such tests. It does not apply to tests developed and administered by colleges for their own purposes. [49:1] The bill states that after the test has been filed with the commissioner of education, and upon request of the test subject, any test taker may obtain, for a nominal fee (1) a copy of the test questions used in determining the raw score, (2) the individual answer sheet together with a copy of the correct answer sheets, and (3) a statement of the raw score used to calculate the scores already sent to the test subject. [287:8]

Under the Weiss bill, test agencies are to report complete testing costs and fees to the commissioner, including the total number of times a test was taken during the testing year; the number of test subjects who took the test once, twice, and more than twice; the total amount of fees received from the test subjects; the total amount of revenue received from each test program; and the expenses incurred by the test agency for each test program, test development, and all overhead expenses. [287:10-11]

Within one year after the bill's enactment, the commissioner is to report to Congress the relationship between test scores and income, race, sex, ethnic status, and handicapped status; and the relationship between test scores and the completion of test preparation courses. [287:7]

Representative Weiss said that his bill did not dictate what questions can or cannot be asked on a test nor did it "try to alter the admissions criteria developed by colleges, graduate and professional schools." [98:11] In testimony before the House Subcommittee on Elementary and Secondary Education, Paul Pottinger, director of the National Center for the Study of Professions, cautioned not to let "clever and professional marketing experts. . . confuse and impress" legislators with technicalities and test development costs. [24:2] Also in testimony before the House Subcommittee on Elementary and Secondary Education, Althea Simmons, director of the Washington Bureau of the NAACP, testified that truth-in-testing legislation was an idea whose time had come. The NAACP was aware that test scores were supplemented by other admissions criteria Simmons said, but College Board scores were still "the deciding factor regarding *who* gets admitted to *what* institution." [256:7] [emphasis in the original] Simmons stated that the NAACP did not oppose use of tests per se, but they did call upon the testing industry to develop standardized tests which have been corrected for possible cultural bias and which properly measure the amount of knowledge retained by students regardless of their background. [259:8] Release of test questions and answers was seen as an overstated concern. [256:8] "The intent of testing," Simmons concluded, "should be to open up avenues for students." [256:11]

Test agencies such as ETS and ACT have vigorously opposed both the Gibbons and Weiss truth-in-testing bills. Test agencies said that these bills would create a cumbersome and unnecessary bureaucracy. Fred Hargadon, chairman of the College Board, urged Congress to let the testing legislation "ferment" until everyone could learn from the test disclosure experience in New York. Hargadon cautioned against federal legislation saying it would result in "regulation of college admissions testing programs, just one step removed from federal regulation of college admissions themselves." [98:11] Clark Cahow, assistant provost at Duke University, said that "federal intervention in the matters of standardized testing when there is no perceived crisis will not change the way schools admit students." [54:2] According to John T. Casteen III, dean of admissions at the University of Virginia, federal legislation would fall short of reforming education because it would sacrifice "the effectiveness of the existing tests as measures of where we are and where we need to go." [54:2] Commenting on criticism leveled at the New York law, Albert Shanker, president of the American Federation of Teachers, called federal legislation "dangerously premature." Shanker expressed concern about preserving the quality of the tests, maintaining standards, and the drastically altered role of the federal government in education which would result from the passage of truth-in-testing legislation. [54:1] Former U.S. Commissioner of Education and president-elect of the Carnegie Foundation for the Advancement of Teaching, Ernest Boyer, rejected the idea that Congress needed "to tamper with testing." Boyer called upon colleges and schools to cooperate more closely to improve the quality of both tests and schools. Boyer stated that the truth-in-testing campaign was diverting the nation from the real issue, quality in education. [30:3]

An August 1979 editorial in the *Washington Post* stated that the Weiss bill attempted to establish broad federal supervision over all admissions testing and helped lay the "foundation for a regulatory system." The editorial commented that test scores are not the only criterion for college admissions and that there was no reason to diminish their usefulness and force institutions to rely on more subjective criteria, not necessarily to the advantage of "the youngster who does not fit the usual pattern." Furthermore, since oversight of the tests would reside in the Department of Education, where the NEA has strong political influence, the implications of the law exceeded the bill's intended purposes, the editorial concluded. [248:A26]

After what was termed a "hysterical lobbying effort" on the part of the testing organizations, the Weiss bill was withdrawn. [215:A4] An aide to Congressman Weiss said that, rather than have the measure pruned of its salient features, further action would be postponed until 1980 when more congressional support could be gathered. [4:6] Don Cameron, assistant executive director of the NEA said that "the testing industry exerted all the influence it has to kill this legislation because they don't want to reveal what this legislation would make them reveal." [215:A4]

Rexford Brown and Merle McClung pointed out that if truth-in-testing legislation survives in its present form, litigation may expand to include other forms of testing, and test takers who challenge the validity of a test or who allege misuse of a particular test may also initiate litigation. Therefore, until the legal parameters are firmly established, they suggested that some states may wish to reconsider implementing truth-in-testing legislation. [34:41-42]

Public Interest Principles

Representatives of five national testing programs established a set of "Public Interest Principles" in January 1980. Leaders of the College Board, Graduate Record Examinations Board, Graduate Management Admission Council, Law School Admission Council, and Educational

Testing Service stated that public interest principles would guide future developments in their standardized admissions examinations. [72:1]

These principles call for increased publication of test content, publically visible methods for eliminating test bias, procedures for test takers to have their scores verified, and ways to increase the appropriate use of scores and to discourage their misuse. However, because individual testing programs differ according to the purpose of the institutions using them, changes will be developed separately by each organization. [72:1] One of the purposes of the guidelines, asserted Hanford, was "to get on record in advance of a number of legislative initiatives outside New York that we are taking seriously the expressions of public concern and intend to do something about it voluntarily." [108:11] "The point is," Turnbull added, "that we are listening to the comments people are making about the need for more public information." [294:137] Senator LaValle noted that the test sponsor's guidelines indicated "a willingness on the part of the testing agencies to comply with the law." [108:11]

The text of the "Public Interest Principles for the Design and Use of Admissions Testing Programs" follow:

1. We recognize the legitimate interest of the public in knowing what the tests contain and their efficacy in performing their intended functions. Therefore, we will implement the principle of publication of test content to a degree limited only by reasonable safeguards of efficiency, cost, quality, and the educational impact of the programs.

2. We fully support the principle of equity and we will continue to maintain and strengthen credible procedures for detecting bias and eliminating it from the content of the tests, while making such procedures visible to the public.

3. We recognize the need for routine procedures that allow the test taker to arrange for verification of the accuracy of the procedures determining the score attributed to him or her.

4. We believe that tests should be readily available to all individuals, regardless of conditions such as physical handicap or religious beliefs that may prevent the taking of exams under circumstances that meet the convenience of the majority.

5. We recognize that tests, together with the procedures for scoring them and reporting the results, should be designed to provide test takers with as much useful information as may be feasible about the specifics of their performance on the tests.

6. We reaffirm the right of individuals and institutions to privacy with regard to information by and about them, which should be safeguarded from unauthorized disclosure.

7. We recognize the need to formulate, maintain, and publish widely principles of appropriate use of scores and other test information derived from testing programs and to be alert to and actively discourage misuse.

8. We recognize that both the institutions making use of test scores and the test takers themselves should have mechanisms through which to express their legitimate interests concerning the design and operation of testing programs and the use of the information derived from them. [73:1]

See Appendix E on page 147 for the complete "Operational Elements" of the Public Interest Principles in which test sponsors present their diverse approaches to opening communications concerning important issues in testing.

While sponsors of the GMAT, LSAT, and GRE have stated they plan to make disclosure their national policy, other test sponsors have indicated that they will continue to oppose truth-in-testing legislation. [108:11] Hanford stated that the steps outlined in the "Public Interest Principles" should not be interpreted as a change in the College Board's opposition to truth-in-testing, which he characterized as "educationally unsound" and a "threat to the validity and fairness of admissions testing." Similarly, in a letter to the editors of the *New York Times*, Turnbull said that ETS continued to oppose testing legislation in New York and elsewhere because the disclosure of every edition of every admissions test has "the effect of curtailing the number of tests that can be administered, raising the costs to students, and threatening test quality." Turnbull asserted that the "Public Interest Principles" call for increased publication of information about tests that goes "beyond any law in these matters." [282:1]

According to one educational newspaper, however, the only real concession contained in the "Public Interest Principles" was that groups of test sponsors and test makers would probably announce when the results from one of their tests would be available. [294:137]

Summary of the Truth-in-Testing Controversy

Test sponsors, agencies, and their supporters assert that truth-in-testing legislation measures are not consumer bills but actually consumer fraud. [71:2; 231:190] The New York State legislation, they believe, has signaled the beginning of government regulation over the admissions testing process, although no need for such legislation has been demonstrated. Many testing organizations find themselves in agreement with those who believe that test takers and the public need a better understanding of the uses and limitations of tests. [71:2-3] But they strongly oppose legislation which mandates the disclosure of every edition of every test. The total disclosure provision threatens test quality, increases costs to students, and reduces accessibility to testing programs by handicapped citizens, Sabbath observers, military personnel, and walk-in applicants. [133:2; 65:2]

In addition, test agencies emphasize that most information required by test legislation regarding test content scoring, score reporting, and interpretation is currently being provided or will be implemented soon. [133:4] The College Board reported that its routine policies and practices voluntarily meet many of the truth-in-testing requirements. CEEB reiterated that it provides a complete practice SAT to all test takers, the correct answers, scoring instructions, a review of the types of questions, and test-taking strategies before they take the test. Additionally, the College Board distributes, to educational institutions and students registering for the SAT, "Guidelines on the Uses of College Board Test Scores and Related Data," which addresses many concerns such as the nature, purpose, and use of the test. [53:2]

Test sponsors and agencies are urging the postponement of any additional laws to regulate admissions testing until the New York experience has been evaluated, legal rulings have been reached, and independent studies completed. [71:2] ETS believes that it is counterproductive to attack standards at a time when the public is calling for quality education, and further, that the anti-testing movement is not supported by the majority of Americans. [71:2] "If the motivations behind test disclosure legislation are sooner or later to put an end to admissions tests altogether," stated George H. Hanford, president of the College Board, "then the admissions process will become more and more capricious." [51:4]

Proponents of truth-in-testing, on the other hand, have argued that standardized tests are a key factor in getting into college and professional school. [36:1] Truth-in-testing legislation, they assert, will open the testing process to increased understanding by the public of the purposes and limitations of tests. [130:5] Supporters of testing legislation state that they are not seeking the elimination of standardized testing but rather encouraging its responsible use. [149:2] They maintain that truth-in-testing legislation is not going to put test makers out of business, reduce their corporate profitability, or do away with college admissions tests. [226:2] Furthermore, since test legislation advocates contend that the cost of test development is quite small, disclosure provisions should not hamper the development of new tests nor increase significantly the fees for test taking. [292:3] They also view testing legislation as promoting competition among test agencies and assuring a degree of quality control through public review. [236:199] As more states follow the example of New York, and federal legislation makes its way through Congress, advocates of truth-in-testing maintain that the mantle of secrecy which surrounds the testing industry will be broken, and light shed on them for the first time. [153:64]

In June 1980, New York Governor Hugh Carey said that most of the legal and administrative difficulties stemming from the enactment of the standardized testing legislation had been remedied by amendments to the law. Carey reported that the testing law had not resulted in the adverse consequences predicted. Overall, there had been no dramatic escalation in the cost of such tests nor a massive reduction in their availability to New Yorkers. [37:1]

Rexford Brown and Merle McClung pointed out that the testimony of proponents and opponents of truth-in-testing legislation abounds with assertions but little substantial evidence. The lack of hard data and strong emotions evoked by the debates, they maintain, are signs of deeper, undeclared issues which have troubled the nation for generations: meritocracy, elitism, and equal opportunity. [34:49] The following exhibits prepared by Brown and McClung and reproduced in Table 18 summarize both the pro- and anti-testing stances regarding truth-in-testing legislation. [34:xi-xiii]

TABLE 18.--The Debate Over Testing Legislation: Pros and Cons

I. Debate About Role and Power of Testing Companies

Anti-Testing Sentiments	Pro-Testing Sentiments
Tests have profound influence upon American lives and life chances.	Tests have little influence compated to family, social and educational influences, grade point average.
Testing companies are unaccountable to their dependent public--particularly true of the Educational Testing Service (ETS) and American College Testing (ACT).	Commercial test publishers are accountable to market forces; test makers, including ETS and ACT, are accountable to professional standards, education community, higher education communities, courts, client groups, trustees, and Internal Revenue Service.
Testing companies are too secretive; test takers do not know about tests; researchers cannot study them.	Critics confuse security--a technical issue--with secrecy; ample test information is available both to test takers and qualified researchers.
Massive testing does more harm than good (*e.g.*, consumes time better spent learning, alters curriculum, stigmatizes children, misleads public, etc.)	The public and higher education have asked for massive testing; testing produces information useful for improving education; it does more good than harm (*e.g.*, takes little time, diagnoses problems, helps administrators, etc.)
Tests are inherently biased against pluralism, tend to further stratify society.	The culture is inherently biased; bias in tests is being minimized; don't blame the messenger; testing helps miniorities.
Tests are widely misused and misunderstood.	Test companies try hard to curb abuse, educate users.

(Continued)

TABLE 18 *(Continued)*

II. Debate About Quality of Standardized Machine-Scored Tests Used Primarily for Prediction

Critics of Standardized Tests	Defenders of Standardized Tests
Tests concentrate on easily measured cognitive skills, ignoring higher level skills (e.g., problem solving), imagination, creativity, etc.	Society values intellectual achievement, cognitive skills; education (especially higher education) stresses those skills; others (e.g., teachers) are better able to assess imagination, creativity, etc.
Theory upon which testing rests is simplistic, outdated and sketchy.	Theory upon which education rests may be simplistic, outdated and sketchy; test theory is better than critics think and always improving.
Test information is much less precise than testers pretend.	Test information is improving in precision and is better than massive subjectivity.
Tests are seldom valid even by test makers' standards.	Many tests are rigorously validated and most do what they are designed to do.
Tests are developed subjectively and always contain controversial items.	Tests are developed by educators and scholars, some of whom always disagree with others; in the main, they do what they are supposed to do.
Test scores do not predict success in later life.	Test scores accurately predict such things as academic success in first year of college, first year of medical or law school, etc.; they are not designed to predict success in later life.
Formal qualities of multiple-choice tests convey messages that undercut reasoning skills, writing ability, accurate perception of the world.	No empirical data are offered to support such fears; testing consumes too little of a student's time to have such effects.

III. Debate About Need for Testing Legislation

Pro-Legislation Sentiments	Antilegislation Sentiments
Grade inflation, misuse have combined to give tests too much influence in admissions decisions.	Higher education's need for students has lessened importance of admissions test scores.
A commitment to "truth in lending," "truth in advertising," sunshine laws and consumerism should extend to an area as important as admissions testing.	Test publishers and higher education institutions already provide ample information and protection; analogies to consumer movements are misleading.
Because admissions tests have such influence, there is an overriding public interest at stake.	There are several competing public interests at stake; critics have not established an overriding need for legislation.
Legislation will promote greater accuracy, validity of tests.	Legislation calling for full disclosure will lower the quality of tests.
Legislation will encourage use of multiple criteria in selection process.	Most institutions already use multiple criteria and test agencies encourage the practice.

(Continued)

TABLE 18 *(Continued)*

III. Debate About Need for Testing Legislation *(Continued)*

Pro-Legislation Sentiments	Antilegislation Sentiments
The admissions test industry is not accountable to anyone.	The industry is accountable to the psychometric profession, market forces, academic community.
	Federal legislation would constitute dangerous, if not unconstitutional, federal incursion into education.
	Legislation interferes with First Amendment right of colleges to determine who they want to teach.

IV. Debate About Full Disclosure and Other Aspects of Legislation

Arguments Against Full Disclosure

Students cannot learn much from examining their test items.

Teachers may try to increase aptitude test scores by teaching the test items, thus damaging curriculum.

Full disclosure will compromise test security; compromised security means less confidence in tests.

Release of items will lead to invalid interpretations and misunderstandings.

Accumulation of disclosures over the years will erode test quality and utility.

Good items are the result of a costly, technical and professional process; they should be husbanded to have long life.

Sample tests are sufficient.

It makes more sense to disclose a full specimen of the test before the test taking session, so test taker knows what to expect.

Disclosure will remove economic competitive incentive to create new and better tests.

If admissions officers lose confidence in test scores, disadvantaged students will suffer.

Disclosure means fewer test administrations per year in order to keep a test secure as long as possible.

Disclosure will increase test development costs, thus the cost of tests to students; poorer students will suffer.

Disclosure will decrease amount of time available for development, leading to greater possibility of biased items creeping into tests.

Disclosure will benefit expensive coaching schools, further hurting poor students.

Disclosure makes comparability measurement more difficult.

Disclosure requirement constitutes seizure of private property without due process and in violation of proprietary rights protected by copyright laws.

Counter Arguments About Full Disclosure

Students can learn about tests and test strategy from examining test questions.

Security need not be an issue; new measurement technology could enable testers to eliminate the problem.

Development costs would not increase as much as testers suggest.

(Continued)

TABLE 18 *(Continued)*

Counter Arguments About Full Disclosure *(Continued)*

Items now available only to expensive coaching schools would be available to everyone, benefiting poor students.

There are many solutions to the comparability problem; the laws do not adversely affect comparability measurement.

The fairness issue takes precedence over technical matters.

Disclosure will help admissions officers as well as students.

Arguments Against Release of All Studies, Evaluations or Statistical Reports Pertaining to a Test

(para. 341, New York; Sec. 4(a)(1)(A), Weiss Bill)

These provisions may interfere with academic and institutional freedom in violation of the First Amendment of the U.S. Constitution.

SOURCE: Rexford Brown and Merle Steven McClung. *Searching for the Truth about "Truth In Testing" Legislation: A Background Report.* Denver, CO: Education Commission of the States, 1980, pp. xi-xiii.

USES OF TEST SCORES BY THE COLLEGES

There is considerable controversy surrounding the appropriate uses and possible misuses of college admissions test scores. Critics of standardized college admissions tests allege that test scores are routinely abused. Citing examples of students who have been denied admission to the college of their choice by as few as five test points, they maintain that test scores are often the sole criterion used by admissions committees in determining who is accepted and who is rejected at postsecondary institutions. Beleaguered admissions committees, critics claim, are tempted to use easily quantified objective indicators such as test scores rather than relying on other subjective indicators of college potential such as a student's previous accomplishments, letters of recommendation, personal interviews, aspirations and goals, and motivation. [258:169] Additionally, they point out that in some cases cutoff scores are set so high that thousands of qualified students are rejected outright so that their credentials are never even reviewed. [93:85] Because they believe college admissions tests are so widely misused, anti-testing forces advocate abandoning them altogether They urge institutions of higher education to follow the examples of Bowdoin College and the University of Wisconsin which have dropped their SAT requirements. [286:A2] When college admissions officers make decisions which have substantial impact on the lives of individuals, subjective measures are better indicators of a student's potential to perform college-level work, testing critics contend. [188:xiv]

College admissions officers, on the other hand, assert that test scores are not the sole criterion used in determining admissions to postsecondary institutions. As professionals they are aware of the limitations of tests and they watch for their possible misuse. While test scores do play an important role, they are almost always supplemented by various subjective measures to provide a profile of a candidate's ability to compete successfully in college. The argument about the SAT, they state, is not about college admissions *per se* but about who is admitted to a few highly selective colleges and universities such as Yale and Stanford. [85:42] Test makers have observed that concern about the use of test scores comes at a time when college enrollments are declining and when there is less pressure to gain admittance to college compared to the early 1970s. [57:11; 44:1] Nevertheless, they are concerned that attempts to abandon admissions testing may result in heavier reliance on subjective measures which may be influenced by social and racial bias. [173:A-19] While admissions tests are not perfect indicators of a student's college potential, test sponsors view them as a democratizing influence which promotes fair and equal access to higher education, especially for minorities and students from less well-known high schools. [47:3]

A recent survey of undergraduate admissions policies, practices, and procedures, conducted by the American Association of Collegiate Registrars and Admissions Officers (AACRAO) and the College Entrance Examination Board has shed considerable light on the importance of admissions test scores in admissions decision making the minimum SAT scores required by postsecondary institutions, and other information related to the college admissions process.

Survey of Undergraduate Admissions Policies, Practices, and Procedures

During the winter and spring of 1978-79, AACRAO/CEEB surveyed 2,623 public and private two-year and four-year colleges and universities. A total of 1,463 institutions (55.8 percent) responded, including 401 public two-year colleges, 81 private two-year colleges, 333 public four-year colleges, and 648 private four-year colleges. [218:1] According to James E. Nelson, vice-president for program research and planning at the CEEB, the survey established "the first substantial baseline of data about college admissions practices." [44:2] The majority of colleges that use admissions test scores most often employ them to indicate whether an applicant may have difficulty in completing the academic program without special assistance; review the scores as part of an overall evaluation of an applicant's credentials; and inform students of the general academic level of the institution by incorporating them into a freshman profile. [44:4-5] Highlights of the AACRAO/CEEB survey follow.

Importance of test scores in admissions.--

According to the CEEB, relatively few colleges rely solely on admissions test scores in the decision making process. As shown in Table 19, less than one percent of private four-year colleges and less than four percent of public four-year colleges identified test scores such as the SAT and ACT as the single most important factor in the admissions process. However, 54 percent of the private four-year colleges and nearly 60 percent of the public four-year colleges considered test scores to be a very important factor. Nearly 40 percent of the private four-year colleges and slightly more than 27 percent of the public four-year colleges reported that test scores were either one of several factors or a minor factor in the admissions decision. [44:3]

Minimum SAT score cutoff points.--

Forty-two percent of responding private four-year colleges and 38 percent of responding public four-year colleges reported having minimum SAT scores. (Table 20) The mean SAT verbal and mathematical scores for those institutions were 754 and 740, respectively. Although 22 percent of the private two-year colleges compared to four percent of the public two-year colleges required minimum SAT scores, the mean SAT verbal and mathematical scores were higher at the public two-year colleges (650) than at the private two-year colleges (617). Thirty percent of all colleges surveyed indicated that they had "minimum standards" for scores on the SAT; mean test scores at these institutions was 740. [44:2, 8]

Freshman applicants accepted.- Private four-year colleges, which many people think of as being very selective, reported accepting nearly 78 percent of all applicants, while public four-year colleges accepted 79 percent of all applicants. (See Table 21 on page 104.) The comparable figures for private and public two-year colleges were 86 percent and 91 percent, respectively. [218:3] James Nelson of the CEEB said, "the problem may be that the public mostly hears about only a few very select, prestigious colleges. In fact, the vast majority of colleges, four-year and two-year, public and private, are not that hard to get into." [44:4]

TABLE 19.--Importance of Admissions Test Scores in Decision Making, by Type of Institution

Test Scores Are	Public Two-Year		Public Four-Year		Private Four-Year		Private Two-Year	
Single Most Important Factor	1.2%*		3.6%		.9%		2.5%	
Very Important Factor	9.0		59.5		54.2		35.8	
One of Several Factors	14.7	20.9	23.1	27.3	35.2	39.5	35.8	49.4
A Minor Factor	6.2		4.2		4.3		13.6	
Do Not Review/ No Response	68.9		9.6		5.4		12.3	

*Percent of all institutions in study group.

SOURCE: "College Board Releases Early Report of College Admissions Study," *News from the College Board*, November 1, 1979, p. 9.

TABLE 20.--Minimum SAT Scores Below Which Applicants Generally Are Not Considered Eligible for Admission, by Type of Institution

	Public Two-Year	Public Four-Year	Private Four-Year	Private Two-Year	All Colleges
Percent with any SAT minimum	4.0%*	37.9%	42.0%	22.2%	29.5%
Minimum SAT V-M					
Mean	650**	740	754	617	740
Median	642	738	751	585	740

*Percent of all institutions in study group of this kind.

**The mean or median for those reporting any minimum SAT score cutoff.

SOURCE: "College Board Releases Early Report of College Admissions Study," *News from the College Board*, November 1, 1979, p. 8.

TABLE 21.-- Percent of Freshman Applicants Accepted and Enrolled, by Type of Institution

Percent of Students	Public Two-Year	Public Four-Year	Private Four-Year	Private Two-Year
Freshmen Applicants Accepted for Admission				
Mean	90.7%	78.6%	77.5%	86.2%
Median	93.7	81.3	81.7	90.9
Freshmen Applicants Actually Enrolled				
Mean	77.3%	53.2%	47.7%	65.5%
Median	81.4	52.0	45.5	68.2
Accepted Freshmen Applicants Actually Enrolled				
Mean	81.3%	65.8%	59.8%	73.5%
Median	84.4	65.4	57.6	74.7

SOURCE: "An Overview of Findings from the College Board-AACRAO Survey of Undergraduate Admissions, Policies, Practices and Procedures." Prepared for the annual meeting of the College Board, October 30, 1979, p. 3.

SAT score expectations.– Approximately 45 percent of private four-year colleges and 44 percent of public four-year colleges reported that higher admissions test scores were expected from applicants in 1970 than 1978. [44:2, 8] The comparable figures for private and public two-year colleges were 48 percent and 34 percent, respectively.

Minimum grade point average cutoff point.--

Nearly 58 percent of private four-year colleges and 43 percent of public four-year colleges indicated that they had a minimum grade point average below which applicants were generally not considered eligible for admission. However, more that three-quarters of the colleges reported that the minimum grade point average was "C" or below. [44:2, 8]

In a 1980 interview with *Educational Leadership*, Fred A. Hargadon, dean of admissions at Stanford University and chairman of the College Board, emphasized that most institutions use cutoff scores "only when a student's grades fall below a certain level." Hargadon cited the California system where admissions personnel check test scores to see if they are high enough to offset poor grades, a procedure which is different from setting a cutoff score without reviewing other criteria. Recent data from a consortium of 30 colleges and universities including Stanford, Harvard, the University of Chicago, and Northwestern showed that in 1978, 430 individuals with SAT verbal scores between 750 and 800 were turned down, but 5,531 individuals with SAT-V scores between 500 and 550 were admitted by those institutions. [5:656]

Credentials required of applicants.--Credentials most frequently required by all institutions were a complete high school transcript or evidence of high school graduation or a General Equivalency Diploma (GED). (See Table 22 on page 106.) Over 92 percent of the private four-year colleges ranked the complete high school transcript first in the credentials required of all applicants, compared to 79 percent of the public four-year colleges. Both private and public four-year colleges ranked admissions test scores such as the SAT and ACT third in importance, while private and public two-year colleges ranked them fourth. Private four-year colleges also placed much more emphasis on subjective measures such as letters of recommendation, interviews, and personal essays than public four-year colleges. [44:3, 10]

Exceptions to formal admissions requirements by selective colleges.--Over 92 percent of the private four-year colleges and 80 percent of the public four-year colleges indicated they had "selective" admissions requirements. (See Table 23 on page 107.) But, 54 percent of the private institutions and 62 percent of the public institutions said they made exceptions to their formal requirements for admission. During fall 1978, seven percent of the freshmen admitted to private colleges and nine percent of the freshmen admitted to public colleges were allowed exceptions to regular admissions requirements. Both private and public four-year colleges reported that exceptions were most often made for adult students, disadvantaged students, and members of racial/ethnic minority groups. [218:17]

Subjective judgments of applicants' personal qualities.--As presented in Table 24, 91 percent of the private four-year colleges and 59 percent of the public four-year colleges ranked student motivation or initiative first in the list of personal qualities that their admissions staffs would consider in determining a student's admissibility. Private four-year colleges ranked evidence of good citizenship or moral character second (83.0 percent) and special skills or abilities third (80.1 percent). Public four-year colleges ranked special skills or abilities second (56.7 percent) and work experience related to intended field of study third (47.5 percent). [218:5]

High school course requirements: 1970 vs. 1978.--

A listing of high school course requirements for admission to postsecondary schools is given in Table 25 on page 109. The AACRAO/CEEB study indicated that course requirements for college admissions have varied little since 1970, with the exception of foreign languages. In 1970 approximately 43 percent of the private four-year colleges and 20 percent of the public four-year colleges reported students having a mean of two-years of foreign language study compared to 38 percent and 15 percent of these colleges, respectively, in 1978. The AACRAO/CEEB survey indicated that college admissions officers seldom rely solely on admissions test scores in determining those who are admitted. [44:3]

According to Rexford Brown of the National Assessment of Educational Progress and Merle Steven McClung of the Education Commission of the States, the AACRAO/CEEB survey data do not prove that institutions are not overestimating the importance of test scores or their accuracy. Brown and McClung contend that additional information is needed about the undergraduate and graduate selection process if the problem of

TABLE 22.--Credentials Required of All Applicants, by Type of Institution

Credential	Public Two-Year		Public Four-Year		Private Four-Year		Private Two-Year	
	Rank*	Percent**	Rank	Percent	Rank	Percent	Rank	Percent
Complete High School Transcript	2	54.7%	2	79.3%	1	92.2%	1	90.8%
Evidence of High School Graduation or GED	1	56.3	1	82.9	2	86.6	2	87.0
Admissions Test Scores such as ACT or SAT	4	18.0	3	65.8	3	74.0	4	56.9
Achievement Test Scores such as ACT, CLEP, Achievement Tests	8	1.2	7	4.3	8	16.5	9	6.0
Letters of Recommendation	7	1.3	6	7.1	5	63.0	5	53.3
Interviews with Admissions Staff, Faculty, Alumni	5	7.3	9	1.3	7	18.1	6	24.5
Personal Essay or Autobiographical Statement	6	2.0	5	10.0	6	44.9	7	15.8
Health Statement	3	33.7	4	54.0	4	63.2	3	73.6
Portfolio, Statement, Audition, etc.	9	1.0	8	2.8	9	10.1	8	8.9

[*] Frequency of reporting requirement of all applicants.

[**] Percent of responding institutions which require credential of all applicants.

SOURCE: "College Board Releases Early Report of College Admissions Study," *News from the College Board*, November 1, 1979, p. 10.

TABLE 23.--Exceptions to the Formal Requirements for Admission, Selective Institutions, by Type of Institution

	Public Two-Year	Public Four-Year	Private Four-Year	Private Two-Year
Percent who are selective	9.4%	80.0%	92.3%	65.4%
Percent of selective institutions who make exceptions to formal requirements	*	62.1%**	53.9%	39.6%
Number of exceptions as a percent of the number of freshmen admitted, fall 1978				
Mean		8.9%	6.9%	4.1%
Median		4.0	4.0	2.0
In 1970, the number of exceptions was				
Higher		15.9%***	20.2%	22.2%
About the same		58.5	59.4	63.9
Lower		25.6	20.4	13.9
In the mid-1980s, the number of exceptions will be				
Higher		17.5%	14.1%	21.1%
About the same		77.0	79.3	76.3
Lower		5.5	6.6	2.9
Exceptions could be made for				
Athletes		43.3%**	24.2%	17.0%
Alumni relatives		24.6	33.4	30.2
Faculty relatives		26.5	36.3	32.1
Members of racial/ethnic minority groups		50.0	41.6	30.2
Disadvantaged students		51.1	38.7	20.8
Physically handicapped students		39.0	29.7	24.5
Students with special talent in art, music, etc.		42.0	30.2	18.9
Adult students		53.8	44.9	47.2

---

*The number of public two-year institutions who are selective is too small for meaningful descriptions.

**Percent of all institutions in the study group.

***Percent of respondents who said they made or could make exceptions.

SOURCE: "An Overview of Findings from the College Board-AACRAO Survey of Undergraduate Admissions, Policies, Practices and Procedures." Prepared for the annual meeting of the College Board, October 30, 1979, p. 17.

TABLE 24.--Percent of Respondents Who Would Consider Subjective Judgments of Personal Qualities, by Type of Institution

Factor	Public Two-Year		Public Four-Year		Private Four-Year		Private Two-Year	
	Rank	Percent	Rank	Percent	Rank	Percent	Rank	Percent
Leadership capabilities	6**	11.3%*	4	44.1%	4	78.0%	4	63.2%
Community or church involvement	7	5.0	7	28.6	6	72.6	6	59.7
Motivation or initiative	2	22.5	1	58.7	1	90.9	2	82.9
Work experience related to intended field of study	1	28.8	3	47.5	7	65.0	7	57.9
Compatability between institutional qualities and student needs	5	14.0	5	40.5	5	77.3	4	63.2
Evidence of good citizenship or moral character	3	21.8	6	37.8	2	83.0	1	83.1
Special skills or abilities	4	19.9	2	56.7	3	80.1	3	65.8

*Percent of respondents who said that factor was sometimes or often important

**Relative frequency with which factor was identified

SOURCE: "An Overview of Findings from the College Board-AACRAO Survey of Undergraduate Admissions, Policies, Practices and Procedures." Prepared for the annual meeting of the College Board, October 30, 1979, p. 5.

TABLE 25.--High School Course Requirements in 1970 and 1978, by Type of Institution

High School Subject	Public Two-Year		Public Four-Year		Private Four-Year		Private Two-Year	
	Any*	Mean**	Any	Mean	Any	Mean	Any	Mean
1970								
English	19.2%	3.7	56.5%	3.6	69.4%	3.8	56.8%	3.8
Mathematics	16.0	2.0	52.9	2.1	67.6	2.3	49.4	2.0
Physical Sciences	12.2	1.6	38.4	1.2	50.3	1.3	30.9	1.2
Biological Sciences	12.5	1.5	43.5	1.1	55.9	1.2	42.0	1.2
Social Studies	14.2	2.5	47.1	2.1	61.3	2.1	46.9	2.0
Foreign Language	3.7	2.9	20.4	2.0	42.9	2.1	18.5	1.9
1978								
English	16.5	3.7	55.6	3.7	71.9	3.8	65.4	3.8
Mathematics	13.2	1.9	51.1	2.1	69.8	2.3	53.1	2.0
Physical Sciences	10.2	1.3	35.7	1.2	51.4	1.3	34.6	1.1
Biological Sciences	10.0	1.3	40.5	1.1	56.8	1.2	45.7	1.1
Social Studies	11.7	2.3	44.7	2.0	63.1	2.1	53.1	2.1
Foreign Language	2.0	2.5	14.7	2.0	37.5	2.1	11.1	1.8

---

[*] Percent of responding institutions which reported a specific number of years' study required

[**] The mean number of years for those reporting any specific requirement

SOURCE: "An Overview of Findings from the College Board-AACRAO Survey of Undergraduate Admissions, Policies, Practices and Procedures." Prepared for the annual meeting of the College Board, October 30, 1979, pp. 6-7.

alleged test abuse is to be solved. In order to ascertain if institutions abuse test scores, they have suggested that it would be helpful to survey admissions officers to estimate their knowledge about the characteristics and limitations of admissions tests. They further suggest that it would be useful to analyze the data that test publishers make available to test score users: "Are their handbooks and explanatory materials readily understandable?" "Do they explicitly emphasize test limitations or are the limitations only implicit in statistics that few readers will attend to or understand?" "Are the handbooks and materials widely available?" "Are they updated often?" "What do admissions people do with the handbook?" [34:52]

Survey of College Admissions Officers

Another survey of college admissions practices focused on the weight of grade point average, rank in class, and other criteria used in determining a student's admissibility to college. In response to a request by the Fairfax County (Virginia) Board of Education, Nancy Kokus, a guidance project specialist, prepared a report on the county's grading system and its implications for students. Included in the 1979 report, *Revision of Grading Scale--Report to School Board*, was a survey of 21 college and university directors of admissions. Each director was asked to comment on the Fairfax County Public Schools (FCPS) grading scale in the selection of applicants. Since there is great variability in the assigned percentages equivalent to letter grades nationwide, Fairfax County wanted to know if its grading practices placed its students at a disadvantage in the admissions process. (FCPS' grading scale is: A=94-100, B=87-93, C=80-86, D=70-79.) Highlights of the responses from the college and university admissions officers follow:

- Amherst College: "We look very carefully at each and every course that the student takes and, of course, how the student fares in each of these courses. What we are most concerned with is the kind of academic program which the student is following, i.e., we are looking for a demanding, rigorous academic program throughout the student's entire high school program, and not necessarily the overall grade point average. [emphasis in the original] An A in an advanced placement course is much more meaningful than an A in a course which is not as challenging."
- Brigham Young University: "At Brigham Young University we use a combination of high school grades and test scores from standardized tests such as the ACT and SAT to determine a student's admissibility. It is difficult to assign a weighting or to describe the weighting that we attach to each because that's put into what is known as a regression formula and the weighting is different each year depending on certain characteristics. The rank in class does provide us important information that we look at in determining scholarships and also admissions for marginal students."
- Duke University: "We base our decision for admission on the total appeal of the student's application in light of its component parts. Special emphasis is given to (1) secondary school academic performance; (2) standardized test scores; (3) extracurricular involvement; and (4) the secondary school's recommendation. We make every effort to evaluate the student's academic record in relation to opportunities available to him in his school setting. In evaluating

the secondary school record, we rely heavily on the explanatory information provided by the secondary school. Our understanding of different grading systems and of varying curricular offerings is to a large extent dependent on profiles and other data provided by the secondary school, such as transcripts which would show variations in grading skills. Class ranks, weighted or unweighted, or decile ratings are helpful in the decision making process. However, we are careful not to jeopardize any student for whom a rank or rating is not presented by carefully considering other indicators in his academic record."

- East Carolina University: "East Carolina University places more weight upon the grade point average or the class rank, as given to us by the secondary school, in determining the admission status of an applicant."
- University of Florida: "The affect of class rank in assessing a student's application is minimal. We do closely weigh, however, an academic grade point average. . . . This office attempts to consider the grading scale when reviewing an applicant's grade point average and admissibility to the University of Florida."
- Harvard-Radcliffe: "Harvard simply takes all the information it can get on every single applicant, from grades and scores, rank in class, extracurricular honors, individual talents and most importantly of all, written observations of an incisive nature from counselors and teachers in order to help us separate students who will look very similar on paper. . . . We do tend to rely very heavily on the subjective data. . .to help split hairs among students who are very, very strong on objective grounds."
- James Madison University (Harrisonburg, Virginia): "Preference is given to those applicants whose program of studies in high school exceeds published minimums. No specific 'weight' is placed on the grade point average or class rank in reaching an admissions decision. Consideration is given to such factors as academic preparation and performance, scores on standardized tests, interests of the applicant, and potential for enhancing the diversity of the University Community."
- Massachusetts Institute of Technology: "The important thing, from our point of view, is that whatever grading scale is used, it allows us to compare the quality of intellectual content in a student's performance in a particular subject with other students. We are interested in using a rank-in-class, for it allows us to compare the students within your school. We have never felt that grade point was a particularly good way of comparing students from one school to another and do not do that. . . . It is important for schools to articulate clearly whatever grading system they use on all transcripts."
- Mount Holyoke College: "There is no question that a student's performance in high school is of the greatest importance in making admission decisions. We, like other colleges, find that the school record is the most important single criteria. We do try, however, to interpret all grades in the light of the particular school, the population, and of course, the grading system itself."
- The University of North Carolina: "As far as weight is concerned, we probably place about two-thirds of our emphasis on the total high school record, including

rank, but also including the method by which the rank was derived, the quality of the academic course selection based on what the school has to offer, and other factors."

- University of Notre Dame: "Grade point averages are not given much 'weight' in our evaluation of students; however, individual grades in specific courses are important in our deliberations. Class rank is an important factor in our consideration. We recommend that schools follow the recommendations of the National Association of Secondary School Principals in determining class rank (weighting grades) when different levels of the same course are offered in the curriculum."
- The Pennsylvania State University: "Our admission decisions are based on an evaluation index which is derived from the high school grade-point average in combination with the verbal score and the math score of the Scholastic Aptitude Test. The high school grade-point average has about sixty percent of the weight in the evaluation index. In order to evaluate applicants as described above, we must have end-of-year grades for all academic subjects. We use the grades exactly as reported by the secondary school. If no grading scale were given, we would ask the school to provide one. While grading scales vary from school to school, grading practices also vary. Therefore it is difficult to say whether a particular grading system places an applicant at a disadvantage."
- Radford College (Radford, Virginia): "The competition for admission may vary in any given year at Radford College. We review the student's application based on the competition within the particular high school that the student attends. Therefore, the grade-point average and rank-in-class are important factors in reviewing a student's application but are not the only items considered in making the decision. Several schools in the State of Virginia do weigh the grade-point average by awarding more credits for advanced placement classes. Other schools have eliminated high school rank which places much higher emphasis on the SAT's. . . . We have rejected many apparently good students from schools who have eliminated class rank where the student's SAT scores have been below the national average."
- University of South Carolina: "The University of South Carolina has a minimum SAT requirement of 350 on each part of the Scholastic Aptitude Test. Many people interpret this to mean a student with 700 may be admitted. . .a fallacious assumption. It is quite possible that a student with 1100 on the SAT will be denied admission. For example, an 800 SAT verbal and a 300 math will not . . . routinely. . .quality the student for admission. On the other hand, a student applying for a humanities major with those scores. . .and a very good class standing would be admitted. . . . We do not normally use raw grade point average for admission, preferring instead to use the more informative rank in class. In the absence of rank, however, we have a formula which allows us to predict academic success based upon grade average."
- Texas A & M University: "The grade point average, of course, determines a student's class rank. This in turn determines the SAT required for admission. At the present time we are requiring nonresident students to rank in the top fourth of their class and score at least 1000 on the SAT."

- University of Virginia: "In our own admission process, we pay more attention to class rank and to the relative difficulty of each applicant's academic program, than to numerical grades. Extraordinarily high or low grades always attract the attention of members of our Committee on Admissions; class rank is of greater importance."
- West Virginia University: "West Virginia University bases the admission decision on the grade point average. Class rank is not considered."
- College of William and Mary: "William and Mary does not have a process whereby we assign 'weights' to specific factors in an applicant's credentials; rather we review the total record of each applicant and compare it with other students applying for admission."
- Yale University: "While grade point average and class rank are certainly considered, they are always assessed in two 'contexts': the context of the student's general academic environment (school system, specific school, course selections, etc.) and the context of the individual application itself--which of course contains much important information which is not academic at all. Thus, we do not assign a standard 'weight' to these factors, but instead try to judge each application, to some extent, on its own terms."

As the letters to Kokus from the admissions officers show, many of the institutions surveyed do not use a specific predictive formula for making admissions decisions. The quality of students' academic program and high school achievement reflected by rank-in-class were of prime importance, with rank-in-class generally considered more important than grade point average. Test scores were used in conjunction with other factors such as academic preparation, class rank, grade point average, student interests, and counselor/teacher recommendations. Several colleges reported that it was beneficial when secondary schools provided a profile that included all possible information about the school and its programs, including the population, the distribution of grades, and the courses offered. As for the FCPS' grading scale, none of the institutions surveyed indicated that it would place the county's students at a disadvantage in the admissions process. [147:11]

An April 1980 article that appeared in the *Washington Post* described the admissions process and the method by which candidates are selected at Georgetown University in Washington, D.C. During the past decade, the number of applicants to Georgetown rose by nearly 80 percent while the chances of being admitted have declined from 57 percent in 1971 to 30 percent in 1980. Of the 7,843 high school seniors applying in the spring of 1980, nearly three-fourths never made it to the admissions committee. These students were either accepted or rejected outright based on a "mathematical index, calculated from class rank, admissions test scores, and the competitive standings of different high schools." From the likely accepts and likely rejects, every application is read by two admissions officers. Hundreds of low-scorers were selected for a second chance and many high-scorers had their scores pulled down because of the type of high school courses taken, the quality of interviews with alumni or admissions officers, or the quality of recommendations from teachers. Alumni relatives did receive some preference in the admissions process. According to admissions director Charles Deacon, "Most of the people in the 'likely reject' group would be able to succeed here, and we want to take some of them. We don't just accept from the top of the list automatically." [93:31, 35]

Although test scores are important in the admissions process, particularly if a student is applying to a selective college or university, it is apparent that other measures such as a student's special talents, accomplishments, experiences, interests and goals are important and may become increasingly important. According to the Educational Testing Service, new ways are needed to insure a fair judgment of those personal qualities. Currently, ETS is in the middle of a five-year project to aid colleges and universities in understanding how information about personal qualities can supplement test scores and grades in the recruitment and retention of students. The Personal Qualities in Admissions project, conducted by ETS in cooperation with nine institutions (Bucknell University, Colgate University, Hartwick College, Kalamazoo College, Kenyon College, Occidental College, Ohio Wesleyan University, the University of Richmond, and Williams College), is studying the academic and personal background of 25,000 applicants at these institutions. It is anticipated that the results of the study will "support a broadened view of admissions." [57:11]

HOW TEST SCORES MIGHT BE IMPROVED

On one end of the testing controversy, anti-testing forces led by Ralph Nader have called for the abolition of admissions tests. [68:2] Other test critics like Allan Nairn have argued that additional criteria should be used to evaluate a candidate's potential to perform college level work, rather than relying on a single testing instrument from a single organization. [188:383, 385] The National Education Association (NEA) has called for free coaching schools for all students, especially for those from lower socioeconomic groups, minorities, and women. [194:26] Testing critics contend that the existence of coachable admissions tests creates financial barriers to educational opportunities, in direct conflict with congressional policies calling for equal educational opportunity. [90:2] In addition, they state that standardized tests have excluded individuals from college, graduate schools, and jobs by as few as five test points. [166]

At the other end of the spectrum, the College Board has asserted that there is no evidence that reasoning abilities "can be taught over a short duration." [121:4] ETS has maintained that institutions using admissions tests are aware that assessments of student characteristics are not infallible, and that the SAT, although not perfect, it is still the best predictor of college success when combined with high school grades. [77:25] Standardized admissions tests were designed to help individuals and institutions at transition points from high school to college, from college to graduate school, and from education to work, but "nobody is kept out of college because of a test score." [123:1; 47:5]

In spite of these opposing points of view between the test agencies and their critics, there are many factors such as students' attitudes and motivation, testing conditions, and familiarization with test taking skills and test content which exert some influence on admissions test scores. Most discussions about how to improve test scores, however, focus on special preparation or coaching.

Perspectives on How Test Scores Might Be Improved

There is no indication that the debate over appropriate methods of improving SAT and ACT test scores will end soon. Many schools are altering their curricula in an attempt to raise their students' SAT scores and it has been estimated that nearly one-third of the public and private schools in the Northeast now offer some type of SAT preparation course. [13:1] This increase of coaching courses in the nation's high schools will, according to the NEA, "further the argument that the SAT is the first course of study that will be taught in the country and, therefore, it provides the precedent for other courses to be used in the development of a national curriculum." [194:31]

ETS, on the other hand, has emphasized that although the SAT is not perfect, test scores are not used alone in reaching admissions decisions. [77:23] The American Association of Collegiate Registrars and Admissions Officers, moreover, reported that less than two percent of selective colleges identified test scores such as the SAT as the single most important factor in the admissions process. [44:1] Additionally, the College Board has stressed that research studies have consistently shown that students cannot increase their test scores significantly by intensive drill. They have stated that research studies, where score increases have been gauged as a result of coaching, showed average increases of under 10 points, less than increases students can expect between the spring of their junior year and winter of their senior year. [50:3] The College Board has expressed concern that the release of test questions and answers would encourage teaching for the test and increase coaching activities. [46:3]

The following letters, official statements, and memos from well-known educators, officials of testing organizations, and critics of admissions tests provide a cross-section of the most recent points of view of their respective organizations and their personal opinions regarding the controversial topic of how test scores on the SAT and ACT might be improved.

Roger Farr, professor of Education at Indiana University and president of the International Reading Association, observed that "although there are contradictory indications, there is some evidence from the NAEP data, SAT scores, and the Indiana study* that any decline that can be supported in reading achievement among U.S. students is one in higher, more sophisticated levels of comprehension. Thus, instead of responding to the alarm over the declining score issue with programs to assure more focus on subskills related to basic literacy, it would seem more appropriate to concentrate on the development of higher level comprehension.

"Because background information has been shown to be a very strong influence on reading comprehension, the surest way to promote higher scores on the verbal sections of tests such as the SAT should be to promote wide independent student reading on a variety of subjects and to expand the background of students on a variety of topics through the use of various media. Since the surest motivation of expansive reading is interest, developing readers who read for their own enjoyment is a prime goal of reading instruction. At the same time, there are aims based on the correlation of reading comprehension to thinking that should guide reading instruction if we are to develop the kind of responses required of students on the verbal section of the SAT.

"A prime goal of teaching reading comprehension should be to enable students to understand, react to, and incorporate or reject *ideas*. [emphasis in the original] To understand, a student must synthesize within and across reading experiences; thus reading should be taught as an induction process that leads to inference drawing as an intelligent guessing game.

"To react, a student must analyze as he or she reads. This involves such thinking habits as challenging generalizations with exceptions and adding options as horns on dilemmas encountered in reading. It means recognizing temporal and causal sequences and comparisons and other symbolism. Such training ought not be reserved for only higher grade levels, but it should be introduced in elementary grades,

---

*According to the National Assessment of Educational Progress (NAEP) data, basic literacy is at a high level. [87:530] The Indiana Study, conducted in 1976, compared the reading achievement of sixth and tenth grade students to their counterparts in 1944-45. Using the same edition of the *Iowa Silent Reading Tests*, it was observed that students in 1976 read as well as or better than students thirty years ago. [86:3] where we tend now to concentrate on only literal comprehension. Inferential comprehension and critical reading begin with learning to use context clues for word recognition, a vocabulary-building practice that depends on the interrelationship of words as ideas; they come too from the practice of predicting the context ahead as one reads.

"To incorporate, a student must learn to test his or her understanding by using what has been learned from reading. This final step of comprehension is too often neglected in classrooms at all levels, where utilization focuses on passing a test.

"It is probable that another way to improve SAT scores would be to drill students on a high-powered vocabulary list, on roots and affixes, and on idiomatic combinations of verb plus acceptable preposition. But the broad reading experience proposed above should develop the same familiarity with word/concepts in a more indelible way; and avoiding the drill will also avoid killing the student's motivation for reading and the separation of language from connected ideas.

"We could free our educational system from this temptation by discontinuing standardized testing that assesses subskills as part of a total reading score. Diagnostic assessment of subskills could continue to guide teacher planning, but we should stop pretending that we know how to prioritize such skills or how they interact in the comprehension process. For the same reason that encouraging lots of reading is the best way to develop readers, summative evaluation of reading by testing should score only reading comprehension as the complex, total act that it is." [155]

Alan C. Purves, principal of University High School at the University of Illinois and president of the National Council of Teachers of English, writes: "College aptitude tests like the SAT-V concern themselves with latinate vocabulary, with subtleties of nuance and tone, and with the comprehension of complex prose. Clearly students can be prepared for such tests through such devices as vocabulary drill, but more importantly through practice in the reading and analysis of a variety of complex material from Shakespeare and T. S. Eliot to essays and other texts by critics, philosophers, and persons of letters generally. Students need practice in close reading and in determining precisely the meaning, tone, and intention of these writers. Such practices need not run counter to other goals of schooling in English, such as the development of aesthetic response, of an expressive writing style, or of a breadth of information about language in general. In fact, a sound English program that does not eschew difficult texts nor a demanding inquiry into the causes of a reader's response would probably enhance a student's aptitude test score. Such a program need not concentrate on drill or other rote studies; these may, in fact, prove counterproductive. Together with careful critical reading, a writing program that emphasizes rhetorical effectiveness for a variety of audiences will enable students to score well on aptitude tests and to achieve good grades in college." [158]

Shirley A. Hill, president of the National Council of Teachers of Mathematics, in a statement of the "NCTM Recommendations for School Mathematics of the 1980s," asserted that "school mathematics is rapidly approaching a crisis stage. Policy makers in education are not confronting the deepest problems because the public and its representatives have been diverted by a fixation on test scores. There are three major urgent problems: (1) school mathematics programs are not keeping pace with the changing needs for mathematical abilities dictated by developing technologies; (2) not enough school time is devoted to mathematics learning and most students are not taking sufficient high school mathematics to prepare them for their futures, as workers, consumers or citizens; (3) there is a growing shortage of qualified mathematics teachers in high school mathematics classrooms." [129:1] NCTM recommended that (1) three years of mathematics in grades 9-12 be required of all students, with diversified programs tailored to the different goals, interests, and abilities of students; and (2) that students planning to go on to college or a vocational program that includes mathematics enroll in mathematics courses during the last year of high school. [129:2] With respect to the evaluation of mathematics programs, NCTM stated that "test scores should not be used as the sole index of success in mathematics programs," and "the results of a test designed for purposes other than program evaluation (such as the SAT) should not be interpreted as an evaluation of mathematics programs." [191:15]

Hill stressed that "these problems cannot be solved by schools, teachers or professional groups alone. The National Council of Teachers of Mathematics offers its resources, its energies, and its volunteer personnel in a massive cooperative effort during this decade. But essential to success is the support and direct involvement of the public and its representatives." [129:3]

Sheila Tobias, educator and author of *Overcoming Math Anxiety*, wrote that very high SAT scores served her well by identifying her intellectual potential and enabling her to be accepted at Radcliffe College. She explained that "it appears that coaching can be effective in preparing students for the mathematics component of the exams more than for the verbal component. . . . There is more of a consensus as to what would be the right answer among those who prepare math problems. Confidence level of course enters in, but unlike the verbal section where, for example, the analogies test not only knowledge of meaning but the cultural context in which the analogy is set, mathematics problem-solving skills can be identified and perfected. For example, students can be trained in how to figure things out: if the problem appears to be far from their previous experience in math, they might substitute simpler numbers, try extreme cases, draw a diagram, and so on. Indeed, to the extent that important skills are being tested (such as problem-solving, extrapolation from what has been taught, etc.) these skills *should* be taught in school. [emphasis in the original]

"What we have learned about 'test anxiety' from years of working with math anxious and avoidant students may not be 'fixable.' They suffer stress when under time pressure; they need to read word problems out loud; they like discussion to stimulate their own thinking. None of these conditions could be incorporated into the testing atmosphere. However, one thing can be learned which is not irrelevant to the kind of thinking that one will need in advanced mathematics: how to go from an answer to a question. . . . Many of us who do well on standardized tests never bother 'doing' a problem at all; we search intelligently among the given answers, discard those that are absurd, concentrate on the two or three that might do, and work pragmatically from the answers back to the question. (Since many wrong answers in mathematics are often right answers to different questions, this strategy also helps to eliminate careless reading of problems.)

"Building confidence is perhaps the most important task in 'math anxiety' or 'test anxiety' reduction. To achieve this we try to get students to become more self-controlled (not give in to panic and stress) and to become more self-instructive. . . . Suffice it to say that helping students find out what is making a problem difficult for them goes far toward achieving both goals.

"Having said this, I do believe coaching can be effective especially in the mathematics area--I am still loathe to recommend it *per se*, both because of an increased focus on test-readying to the exclusion of other important curricular demands, and because I would like to hang on to my old notion that the SAT exam is 'fair.' [emphasis in the original] On the other hand, a correct analysis of what are the skills needed to do well on those exams and training in the use of those skills (irrespective of the exams) might be worth doing in high school." [163]

Vito Perrone, dean of the Center for Teaching and Learning at the University of North Dakota, commented that "secondary schools are responsible for assuring that young people have opportunities to pursue successfully a broad range of academic, social-cultural and vocational learning. If the schools carry out such a responsibility well, most students who wish to go on to colleges and universities, as well as vocational-technical schools, will be adequately prepared. Will these students also score well on college admissions tests such as the SAT and ACT? They *might*, but assuring that students score well on these tests ought not be a principal objective of secondary schools. [emphasis in the original] For students seeking entry into the large number of public colleges and universities, SAT and ACT scores are not particularly critical. Only in the highly selective private colleges, which serve a very small percentage of the college-university population, are the test scores matters of consequence.

"What could secondary schools do if they wished to help students score better on college level admissions tests or if they wished, as some apparently do, to enhance their public image as schools where students do well on admissions tests. In addition to providing for students the highest quality educational experiences possible, which is not in itself likely to guarantee that high SAT or ACT scores will accrue, they could, beginning in the junior year, organize a systematic test preparation program aimed at the SAT and ACT. Such a test preparation program would likely focus on instruction in test-taking--the SAT and ACT tests provide logic systems and formats which are not characteristic of what students typically encounter in their school studies--and would involve experience with past SAT and ACT tests, now more available as a result of the New York Truth-in-Testing legislation. This kind of assistance is offered now by a number of commercial enterprises which have begun to turn large profits from the ability to pay of middle and upper class parents. It is also offered in many of the nation's private schools and academies and is becoming more common in public schools serving middle and upper class populations who prefer entry into selective colleges and universities. Students receiving such test-taking instruction, and clearly the quality of these preparation programs differ, come to the testing process more confident and, in many cases, perform better than might otherwise be the case.

"If the tests are viewed as important for enlarging student's access to post-secondary educational opportunities, then schools should offer specific preparation opportunities." [157]

Willard H. McGuire, president of the National Education Association stated that "research and firsthand reports reveal that being coached, being rich and having access to past test errors will improve high school SAT scores.

"A National Education Association investigation and a Federal Trade Commission study show that coaching for college entrance exams can help to improve SAT scores. A Ralph Nader publication, *The Reign of ETS*, discloses the relationship of high socio-economic status to high SAT scores. New laws provide for disclosure of test results, but proponents of the SAT disapprove and equate the practice with "cheating."

"Since high schools cannot become commercial coaching centers, nor can they cause more openness in testing, nor can they make the poor wealthy, then they must ask other questions and pursue other ways of assessing students.

"What are other ways to assess a student's achievements, strengths, and future directions and ambitions? Some alternate ways are:

"Place greater reliance on a long-time assessment of a student by several professionals at the high school level. Descriptive statements should include a view of academic achievement, strengths, personal qualities, activities, etc.

"Emphasize the value of students' products in the fields of art, science, or the performing arts.

"Assess the value of work experiences and personal events for the student.

"Encourage universities to hold full and in-depth discussions with interested students.

"Seek to open admissions to universities, or establish more liberal policies for receiving students who are willing to accept a challenge.

"A goal for education in the 80's is 'equity' in access for all Americans. Pursuit of higher SAT scores for school admissions is not the road to educational equity." [156]

The American Federation of Teachers stated that "college admissions tests, such as the SAT and ACT, are measures of developed academic abilities. There has been a great deal of controversy recently over whether scores on such tests can be substantially increased through test coaching. Unfortunately, many of the articles which have appeared on this subject have been vituperative, one-sided polemics concerned more with the politics of testing than research findings. Part of the problem has been imprecise use of the term coaching to describe many different instructional activities.

"The American Federation of Teachers believes that short-term drill and practice produces minimal score changes that are more realistically explained through maturation and normal score variations. So called coaching activities which are of sufficient quality and duration to produce meaningful score changes are more accurately described as instruction. Such coaching or out-of-school instruction most assuredly does and should result in greater performance on a test of *developed* academic abilities. [emphasis in the original] The AFT hold that a sound, comprehensive education, equally available to all students, will far better serve students' chances for admission to and success in post-secondary programs than all the coaching courses which may be available." [164]

Scott D. Thomson, executive director of the National Association of Secondary School Principals observed that "the Scholastic Aptitude Test is a measure of verbal and mathematical capabilities important to getting good grades in college. The mathematics section of the SAT examines reasoning, algebraic concepts, and geometric concepts. The verbal section examines antonyms, analogies, sentence completion, and reading comprehension. The 'practice value' of the test is about 15 points, i.e., students who take the PSAT or the SAT will, on the average, increase their scores at the second testing by 15 points.

"High school grades continue to be the best single predictor of college performance. When these school grades are combined with SAT scores, however, a more accurate prediction results than when the grades are used alone. Even when employed by themselves, the predictive validity of SAT scores is substantial.

"While usefully predictive of college performance, the SATs do not measure a number of additional aptitudes related to student achievement. For instance, the important intellectual processes of fluency and originality are not examined. Other qualities of a more affective nature, such as social perceptivity, also are not a part of the SAT instrument. These factors together with student personality traits cause school grades to be more predictive than are the test scores of the current college aptitude tests.

"Recommendations for improving SAT scores are offered within this context. They are not proposed as a panacea because no easy formula exists to what can only be the outcomes of serious learning over a period of time. They are not offered, either, for the single purpose of a high numerical test score. Rather, they are offered for the importance of what the scores represent--competence with verbal and mathematical concepts.

"English: College-bound students should enroll in an English class during each year of attendance in a secondary school. The English curriculum must emphasize, but not be restricted to, expository writing, language usage, vocabulary, and serious literature. Reading should be extensive as well as intensive.

"Mathematics: The successful completion of courses in algebra and geometry, as a minimum, is necessary to good SAT scores. A geometry course is helpful because it introduces the student to deductive reasoning, useful to problem resolution in many walks of life as well as with the SAT test questions. Geometry also provides experience in visualizing three-dimensional figures and in applying formulae for areas and volumes. Advanced mathematics courses, especially Algebra II, trigonometry, and functions (also called precalculus) increase the likelihood of good SAT scores because these courses maintain and reinforce important applications studied only briefly in algebra and geometry.

"Ability Grouping: Students should be ability grouped in English and mathematics as a minimum. Ability grouping should be organized subject by subject according to the aptitude and prior achievement of students. Schools should not 'track' students into broad programs such as 'college prep,' 'vocational,' 'general,' etc. Such tracking practices are unnecessarily restrictive since a student may have a high verbal aptitude and an average mathematics aptitude, or vice versa.

"Special Preparation: Special test preparation opportunities should be available to students on a voluntary basis. Two benefits accrue from a voluntary program of test preparation: (1) individual students with a strong need or interest may attend without obligating an entire class, (2) pressures to base the college prep English or mathematics courses narrowly upon SAT content can be resisted more easily. Test preparation programs seldom exceed a semester and often only consist of 10 to 15 meetings. Whatever their nature, students should have the privilege of participating in a school-sponsored program and not be forced to rely upon the open market for this kind of important professional assistance. Sophistication in test taking, as well as familiarity with test content, can raise student scores. Practice tests, therefore, should be available to students in the counseling departments of schools as well as being a part of special test preparation courses.

"Students and Parents: Student and parental attitudes are critical to good SAT scores. No school curriculum, however brilliantly conceived and taught, can compensate for student disinterest. Ambivalence is the dry rot of a solid academic performance. A strong commitment to excellence should be held by parents, students, and teachers. Schools with stable or rising test scores credit students and parents with high academic aspirations and a commitment to work toward those aspirations." [162]

Diane Ravitch, associate professor of History and Education, Teachers College, Columbia University expressed concern that the subject of "how to improve test scores" represented a confusion of symptom and problem. Ravitch explained, "the problem is not 'how to raise test scores,' but how to improve the ability of students to read, write and compute, and beyond that, how to restore some vision of what it means to be an educated person. To the extent that declining test scores reveal declining competence in the mastery of the basic skills, then we should be alert to the danger signals. To disparage concerns about declining test scores is inappropriate. We would think a medical doctor odd if he chose to ignore thermometer readings of 104°; we would also think him odd if he claimed that his major concern as a doctor was to keep all his patients at 98.6°. We would wonder why he was not asking questions about how to improve their health, rather than focusing narrowly on how to deal with the symptoms of their ill health. To stretch the analogy a bit further, we would also have cause for alarm if the AMA (like the NEA) urged us never to use thermometers at all, or if consumer experts like Ralph Nader) demanded that all thermometers be banned or broken in order not to differentiate between those with fever and those without.

"Tests should be used as a teaching tool, to help the teacher in gauging how well he or she has taught and to help the student in testing his mastery of the material to be learned. If we ask how to improve the teaching and learning of the essential tools of reading, writing, and computing (rather than how to raise test scores), then the answers should be fairly obvious. First, schools should hire the best, most dedicated teachers to be found, should evaluate their performance in the classroom, and should provide them with assistance and opportunities for professional development. Second, teachers should give students the time, encouragement, and reinforcement they need to become skilled in reading, writing, and computing. Third, schools should establish requirements for a common curriculum, so that all students have an understanding of history, literature, social science, foreign language, art, music, mathematics, and science. Fourth, the staff should work with parents to set limits on television-watching and to promote good study habits at home. Fifth, the staff and the community should cooperate to secure an atmosphere in the schools that is conducive to teaching and learning, orderly, free of political pressures, and assured of adequate public support for education. If all of these conditions were met, test scores would not be an issue." [159]

Philip R. Rever, director of the Washington, D.C., office of the American College Testing Program, stated that the ACT's historic position has been that "the results of the Assessment are not uniformly valid indices of secondary school effectiveness." Rever emphasized that "it would not be appropriate for ACT to guide or try to influence school officials' decisions about instruction and curricula. Moreover, as indicated in our statement regarding 'coaching,' it is not clear that schools can or are able to improve students' scores. Even if it were possible, a larger question remains--should schools try to improve students' scores on admission tests?" [160]

The official "ACT Statement on Coaching" (October 1979) emphasized that "the ACT tests are designed to measure the knowledge and skills that students have developed over a period of many years; that is, they are tests of general educational development. As such, they are intended to indicate a student's level of development in the areas tested only at a particular point in time. Growth in a student's capabilities will be reflected in higher test scores, a fact already acknowledged by ACT in the adjustments it makes in test scores for probable growth as a function of the time in the student's education when the test is taken.

"All four of the ACT tests are curriculum-based. The English Usage and Mathematics Usage tests particularly, emphasize the application of content specific knowledge and skills. The Social Studies Reading and Natural Science Reading tests, while also requiring content knowledge, emphasize proficiency in interpretive and analytic skills.

"The ACT tests are used for several purposes including, in some instances, selection decisions related either to admission to undergraduate college or to specialty programs within a college. The primary uses of the tests, however, are in academic advising, career counseling, placement, and institutional research.

"Because the tests are used mainly for non-exclusionary purposes. ACT advises examinees not to engage in intensive preparation for the tests. Careful study by students of the materials provided by ACT at the time of test registration is considered sufficient for familiarizing most examinees with test type and format.

"Because of the curriculum-based character of the tests, particularly the English and mathematics tests, it is possible that, for some students, a thorough review of the discipline sampled by the tests could increase their knowledge and hence, their scores on the tests. This is most likely to occur in those instances where the students have previously acquired the knowledge and skills to be tested but have used them infrequently and therefore have not maintained them.

"Intensive study in the disciplines tested will result in increased proficiency in the knowledge and skills measured by the ACT tests. Improvements in students' test performances and in their test scores will follow. Indeed a necessary attribute of the tests is that they be sensitive to instruction. Such instruction is comparable to that provided through the regular school curricula, the latter being sufficient if the students approach their studies in a serious fashion." [17:3]

George H. Hanford, president of the College Board, in a June 1980 memo to College Board members, committees, and councils, stated that "since SAT questions call for learned responses, and do not measure some innate capacity of an individual, the central issue in special preparation for the SAT is how these responses are learned.

"Foremost among the developed abilities tapped by the SAT is the ability to read with understanding and to reason with words and numbers. One cannot point to any particular part of the school experience where these are taught to the exclusion of the rest of the curriculum. In fact, they are exercised and developed through their application to course work in a wide variety of academic areas *and* through experience outside the classroom. [emphasis in the original]

"Instructional programs designed to improve performance on the SAT vary along two important dimensions--the duration of the instruction and the content of the program. Common sense would lead one to expect more improvement as more time and effort is spent on preparation. Similarly, one would expect greater improvement as the content of the program moves from drill and cramming on practice exercises (the traditional meaning of 'coaching') to formal instruction in reading comprehension and verbal and mathematical concepts. The general results of research studies on special preparation for the SAT confirm these expectations.

"Many of the studies on special preparation for the SAT were sponsored by the College Board. These studies paid particular attention to the question of whether 'coaching,' meaning short-term drill on practice questions such as those contained in the SAT, improved performance on the SAT. These studies have repeatedly shown that resultant score gains are small, except in the case of certain groups of students for whom review of mathematics was found to be beneficial.

"However, preparation for taking the SAT can obviously be more extensive than such short-term pre-test drill. In some instances programs extend over eight to ten weeks with homework assignments for participants. Some schools have provided special courses for their students lasting a semester or more. These more extensive programs are not so much drill exercises as a focused and sometimes intensive educational treatment. To the extent that these programs strengthen the capabilities of students to perform college-level work, they may represent beneficial educational experiences, which may be reflected in SAT scores.

"What is not known is the degree of benefit that can be expected by a given individual under what circumstances, and whether such benefit, if realized, would represent an improvement over regular study with equivalent motivation of academic subjects in school. All research relating to this topic to date has been subject to one or more design limitations whose effect on the outcome remains unknown. Some studies have fewer such flaws than others, but none has been entirely free of them.

"The College Board believes that the following statements with respect to preparation for the SAT are reasonable interpretations of the evidence now available:

(1) The SAT measures developed verbal and mathematical reasoning abilities that are relevant to success in college-level work; it is not a test of immutable individual endowment.

(2) Scores on the SAT are accordingly subject to improvement as educational experience, both formal and informal, causes these verbal and mathematical abilities to develop.

(3) Development of the abilities measured by the SAT is related to the time and effort spent; *short-term drill* is likely to have little effect; well designed, *longer preparation* can have greater effect. [emphasis in the original]

(4) While drill and practice on sample questions show little score increase, such preparation familiarizes the student with different question formats and may reduce apprehension about what to expect; students can also help themselves to become familiar with test format and taking the test by using the full form of the practice test (*Taking the SAT*) with accompanying information provided by the College Board at no charge to each test registrant.

(5) Whether longer preparation, apart from that available to students within their regular high school curriculum, is worth the time, effort, and money is a decision that individuals and schools must make for themselves; effectiveness seems to vary considerably from program to program and for individuals within any one program; studies of special preparation programs available in many high schools show a range of effects averaging about 10 points in Verbal and 15 points in Math. Recent studies of commercial coaching show a wide variability, from negligible effects to average gains (in the case of one school) in the 20-30 point range for Math to perhaps half that for the Verbal. Students and schools should seek evidence that the results of a given program or course are likely to make a difference in relation to their college admissions objectives.

"In general, the soundest preparation for the SAT, in terms of an organized educational experience, is a broad academic secondary school curriculum with wide-ranging outside reading; in any case, SAT score increases of 20 to 30 points correspond to approximately three additional items correct; this effect may as well or better be obtained if students pursue independent study in addition to regular academic courses." [122:3-4]

A Continuing Controversy: Future Role of Test Scores

From its inception 80 years ago, the College Board's stated commitment has been to expand educational opportunity and open access to higher education through standardized testing. [51:3] The accuracy, objectivity, and comparability of these tests have served to identify otherwise undetected talent and enfranchise individuals into the mainstream. [69:4] During the past two decades even the most selective institutions such as Harvard, Princeton, and Stanford have diversified their pools of applicants at the undergraduate and graduate levels. The critics of admissions tests, according to Dennis Gray of the Council for Basic Education, are primarily interested in the relationship between testing and power. The leaders of teachers' unions are trying to "aggrandize teacher power" and thus ensure that test results cannot be used to question the competency of their membership. According to Gray, testing critics such as Ralph Nader believe that the abolition of admissions tests will open up college admissions and eventually corporate power will be effectively managed by better individuals. [117:8]

The continuing controversy over the fairness of standardized college admissions tests is not likely to subside. In June 1980, a Congressional truth-in-testing measure was revived for hearings amid discussions about the meaning of the word 'aptitude' and whether coaching improves SAT scores. [280:2] The National Institute of Education, according to former director Michael Timpane, is preparing to study the impact of testing on college admissions on a federal level to ensure that testing is equitable and leads to fair educational decisions. [212:4] George Hanford has acknowledged that the importance of the SAT may be on the wane, and has said that the College Board is now preparing for a time when testing tools will be "useful in earlier guidance, in placement and assissment for remedial pruposes." [123:5] Fred M. Hechinger, of the *New York Times* observed that the SAT may become "an academic paper tiger within a few years" because colleges are hurting from overexpansion and declining enrollments. He predicted that the purpose of admissions tests "will shift from selection to placement on the campus," but that admissions tests will be subject to attack as long as test critics attempt to substitute "objective judgement for consumers' desire." [127:8] Lastly, according to Andrew Strenio of the Huron Institute for the Nation Consortium on Testing, "if the testing debate bears no other fruit, perhaps it will at least prompt closer consideration of the exact functions of these standardized testing programs and the rights and interests of those tested under them." [266:13]

SUMMARY AND CONCLUSIONS

The current national controversy over standardized testing affects American education from elementary schools to institutions of higher education and confronts educators, school officials and others with difficult questions. For nearly two decades many educators and much of the public have been concerned and alarmed by reports of declining college admissions test scores. Broad media attention focusing on declining college admissions test scores has resulted in public concern and a demand for immediate remedies for what many perceive to be a decline in educational quality throughout the nation.

College admissions tests are designed to measure abilities and skills developed over many years, both in and out of the classroom. Although standardized tests such as the SAT and ACT were not designed to indicate elementary or secondary school effectiveness, evaluate school instructional programs, or indicate an individual's innate intelligence, such functions have become popularly accepted as attributes of these tests. Initially these tests were developed to supplement the college applicant's previous academic record and accomplishments; place applicants from diverse economic, geographic, and educational backgrounds on an equal footing; and assist in academic counseling and placement. Although college admissions tests continue to serve as admissions tools, they have gained public visibility and interest that is not in keeping with their intended function. There are many who believe that the problem is not specifically how to improve the test scores of college-bound students, but rather how to improve their skills along with the skills of other students to read, write, compute, and comprehend at higher levels.

The relationship between declining test scores and possible causal factors, and the relationship between special preparation or coaching and improved admissions test scores, are highly complex, with many variables. In examining college admissions testing and test score trends, school officials and other concerned persons face questions such as: What do changes in admissions test scores mean for our students, school, and community? How can we ensure that we are providing our students with the highest quality educational experiences possible? Should special preparation for college admission be a major focus of our secondary schools? Which types of students might benefit the most from any such special preparation? How can we ensure all students equal opportunity of access to the educational institutions of their choice? Are there better ways of assessing student potential for college admission?

It is difficult to separate the political aspects of the standardized testing controversy from the tests themselves and focus on practical educational decisions that provide the basis for sound education. Additional research is needed to determine exactly the effects of both short-term and long-term preparation or coaching on student learning, particular student populations, circumstances under which an individual may benefit from coaching, whether coaching is better than regular classroom instruction with equivalent motivation, and how this may translate into improved admissions test scores. Additional data are not currently available, so important educational decisions and policy evaluations must be made based on the information at hand. To date this information tends to support the following *tentative* conclusions for consideration when school officials formulate educational policies concerning changes in admissions test scores, the preparation of college-bound students for college admissions tests, how test scores might be improved, and related aspects of standardized testing.

Trends in Standardized College Admissions Test Scores

- SAT test scores for males and females declined between 1967 and 1980. During this 13-year period, the combined male and female SAT-Verbal scores declined 42 test points out of a total range of 600 test score points (from 200 to 800). SAT-Mathematical scores declined 26 points out of a total range of 600 test points. These declines appear somewhat different when interpreted in terms of the number of test questions involved. Compared to their 1967 counterparts, SAT test takers in 1980 answered approximately 6 fewer SAT-V questions correctly out of a total of 85 questions, or 7.0 percent fewer questions. On the SAT-M, 1980 test takers answered 2.88 fewer mathematical questions correctly out of 60 test questions, or 4.8 percent fewer. These two declines combined represent a total annual decrease of two-thirds of a test question for the average SAT test taker over the 13-year period.
- Between 1967 and 1980 the SAT-V scores of males declined 35 points (7.6 percent) and their SAT-M scores declined 23 points (4.5 percent). These declines are equivalent to approximately 5 fewer SAT-V questions and 2.55 fewer SAT-M questions answered correctly by male test takers in 1980, compared to their counterparts in 1967.
- Between 1967 and 1980, female SAT-V scores declined by twice as much as female SAT-M scores, with SAT-V scores of females declining 48 points (10.3 percent) and their SAT-M scores declining 24 points (5.1 percent). These declines correspond to approximately 6.85 fewer SAT-V questions and 2.66 fewer SAT-M questions answered correctly by female test takers in 1980, compared to their counterparts in 1967.
- Over the past six years, the percentage of female SAT test takers increased steadily to an all-time high of 51.8 percent in 1980, while the percentage of male SAT test takers experienced a corresponding decline to 48.2 percent.
- The percentage of ethnic minority students taking the SAT has increased from 15 percent in 1976 to a record high of 18 percent in 1980. Women comprise 55 percent of all minority test takers.
- A comparison of SAT test score averages for minorities shows that between 1972-73 and 1976-77 blacks averaged 119 SAT-V points and 134 SAT-M points below whites, and Chicanos averaged 80 SAT-V points and 81 SAT-M points below whites [135]. Test makers claim that these differences are not due to bias in the tests but reflect the disparity in educational opportunity for minority youths. They contend that, "The tests do not create the inequality; they reveal it." [68:3]
- When making comparisons between two students on the same test, the College Board has stated that differences of less than 1.5 times the standard error of the difference should not be considered significant. In other words, the scores of test takers *must differ by more than* the following number of

test score points before one can conclude with reasonable certainty that the higher score of one test taker actually represents greater ability or higher level of achievement than the lower score of another test taker as measured by these tests: SAT-V, 65 points; SAT-M, 69 points (on the College Board's 200-to-800 scale); TSWE, 7.8 (on a 20-to-80 scale); American History and Social Studies, 68 points (on a 200-to-800 scale for all achievement tests); Biology, 63 points; Chemistry, 59 points; English Composition, 71 points; English Composition with Essay, 69-81 points; European History and World Cultures, 66 points; French, 60 points; German, 57 points; Hebrew, 42 points; Latin, 77 points; Literature, 80 points; Math Level I, 75 points; Math Level II, 72 points; Physics, 62 points; Russian, 60 points; and Spanish, 56 points.

- Because of the large number of points attributable to the standard error of the difference when comparing the scores of two applicants (65 SAT-V points and 69 SAT-M points) a college with arbitrary cutoff scores may be unintentionally excluding qualified applicants of equal ability from consideration on the basis of test scores alone. For example, a college with a cutoff score of 670 on the SAT-V and 680 on the SAT-M may consider the credentials of an applicant meeting these initial requirements, but reject an applicant with scores of 605 on the SAT-V and 611 on the SAT-M, although there may be no essential difference in the abilities of the two applicants, measured by their SAT scores. The College Board advises both high school counselors and college admissions officers "against making fine distinctions between scores."
- Colleges that have arbitrary cutoff scores on the SAT, below which an applicant's credentials may not be reviewed, might consider establishing broad ranges of acceptable scores in order to ensure that some minority students, underachievers, or persons who may not "test out" well on standardized tests are not rejected from consideration based on test scores alone.
- Because of the large standard errors associated with the SAT and achievement tests, caution should be taken not to assume that small differences in scores represent a true difference in test takers' abilities or to assume that test takers can be ranked in terms of their abilities on the basis of small differences in their test scores. At best only substantial differences in test scores should be interpreted as significant indications of true differences in abilities or achievement as measured by the test.
- Between 1975 and 1980, the average score for both males and females combined on the Test of Standard Written English declined by 0.8 points from 43.2 to 42.4. Each of the 50 test questions is worth approximately 1.0 point on the College Board's scale of 20-to-80. Thus, this decrease over the past five years corresponds to slightly less than one test question.
- Unlike the continuing decline in SAT test scores, achievement tests present a mixed picture of some test score declines, some stability, and some improvements. The Math Level I and Math Level II tests have shown a slight downward trend in the late 1970s and into 1980. The European History and World Cultures test, Spanish, and Latin tests have exhibited somewhat erratic trends. With the exception of a one-year aberration in test scores in the English Composition, the American History and Social Studies, and Physics tests, these scores have been rather stable. Likewise, the Chemistry, Literature, French, and German tests have shown overall stability. Test scores in Biology and Hebrew

have exhibited an upward drift in the late 1970s and into 1980, while Russian test scores have climbed dramatically.

- Despite broad yearly fluctuations, when viewed from a perspective over the entire history of the Preliminary Scholastic Aptitude Test/National Merit Scholarship Corporation testing program (from 1959-60 to 1979-80), total PSAT-V scores for males and females combined have declined 1.0 test point from 41.2 to 40.2 or 2.4 percent. During this same period, the combined total PSAT-M scores for males and females have increased 0.3 test point, rising from 45.0 to 45.3, or 0.7 percent. Thus, over a period of two decades there has been a decline of less than one percent (0.8 percent) in the average combined PSAT verbal and mathematics score for all persons taking the tests. Each of the 65 PSAT-V questions and each of the 50 PSAT-M questions is worth approximately 0.9 point and 1.0 point, respectively, on the College Boards 20-to-80 test score scale which corresponds to the SAT's scale of 200-to-800.
- The American College Testing Program's Assessment Program is designed to require students demonstrate reasoning abilities, problem solving skills, and knowledge in four areas: English usage, mathematics usage, social studies reading, and natural science reading. The ACT has a composite score for the test ranging from 1 to 35 test score points.
- Mean Composite ACT scores for a 10 percent sample of males and females combined who took the test increased one-tenth of a standard score point each year between 1975-76 and 1978-79. This increase over the three-year period was noteworthy, according to ACT, because only once since 1969-70 had there been an increase in the average Composite score. The 1979-80 ACT mean Composite score declined one-tenth of a standard score point to 18.5, the same level recorded in 1977-78.
- From 1974-75 to 1979-80, ACT scores for both males and females rose in English, social studies, natural science, and on the Composite score. Mathematics scores of males declined over this period, from 19.2 to 18.9, while females' scores increased from 16.0 to 16.2.
- According to the ACT Assessment Program, a composite score below 14 indicates that a student has had a restricted educational development background; scores between 14 and 19 are considered low average; between 19 and 24, high average; and above 24, superior. The typical student in the 1979-80 sample had an ACT Composite score of 18.9 and a high school grade point average of 3.0.
- Regarding racial composition of the ACT sample, 75 percent identified themselves as Caucasian American; 7 percent, Afro-American-Black; and 2 percent, Mexican American/Chicano.

Causes of Declines in Admissions Test Scores

- Theories for the decline in college admissions test scores include a host of possible causal factors such as the teaching of minimum skills at the expense of developing more complex skills; a decline in intellectual standards; changes in the promotion policies of schools; changes in the composition of the test-taking population; decreasing amounts of time spent on task; increased drug usage among high school students; and exposure to radioactive fallout from atomic bomb blasts. The extent of influence of some or all of these elements, if any, remains vague.

- The "blue ribbon" Advisory Panel on the SAT score decline, appointed by the President of the College Entrance Examination Board in 1975, concluded that there were actually *two* separate SAT score declines characterized by different causal factors.
- The first decline between 1963 and 1970 was due primarily to changes in the population taking the SAT. Compared to the past, the SAT was measuring a broader cross section of American youth which included larger proportions of characteristically lower scoring groups of students from disadvanted socioeconomic backgrounds, ethnic minorities, and women. The Panel estimated that this accounted for two-thirds to three-quarters of the SAT score decline during the period.
- The second decline, according to the Advisory Panel, which occurred between 1970 and 1975, was due to pervasive changes affecting both higher- and lower-scoring groups alike and related to events in the schools and the society at large. The Panel concluded that there was no one cause for the SAT score decline, but rather, a "virtually seamless web of causal connections."
- The causal connections identified by the Panel included: reduced continuity of study in major fields; increases in student absenteeism; grade inflation (indicating declining educational standards); increased use of pictures, wider margins, and shorter words and sentences in textbooks; extensive time that students watched television which detracted from homework and time spent developing academic skills; changes in the role of the family, especially the increasing number of children in single-parent families; changing life styles, values, higher mobility, drugs, and increasing problems of discipline.
- The Panel was unsure about how the disruption of national life during the 1967-75 period (Vietnam War, political assassinations, burning cities, corrupt political leadership) might have affected the motivations of test takers. But the Panel noted an apparent dimunition in students' learning motivation.
- The Advisory Panel continued that the SAT score decline was a complex subject filled with nuances, qualifications, and doubts. The SAT is a "limited instrument" and should not be viewed as "the sole thermometer for measuring the health of school, family, and student."
- Although the Advisory Panel's report has been criticized and SAT scores have continued to decline since the report was issued in 1977, the College Entrance Examination Board has not to date published any additional explanations or possible reasons for the continued decline in SAT scores.

The Controversy over Standardized Testing

- What was once largely an internal debate within the psychometric profession concerning the design and appropriate uses of college admissions tests has escalated into a national debate in which test makers are under extreme criticism, if not an all-out attack, from some consumer groups, public interest groups, and educator organizations.
- The economic and political overtones of the testing debate, the gravity of the assertions and counter-assertions, and the intense media attention, make it difficult to assess the true merits of arguments by both the test critics and test makers.
- Among the prominent individuals and organizations generally critical of standardized

testing and the role of the test makers are Ralph Nader and Allan Nairn, consumer advocates; Willard H. McGuire, president of the National Education Association; Benjamin L. Hooks, president of the National Association for the Advancement of Colored People; Hugh L. Carey, Governor of the State of New York; Kenneth P. LaValle, New York State Senator (R); Steve Solomon, New York Public Interest Research Group; Lewis W. Pike, National Institute of Education; Warner V. Slack and Douglas Porter, Harvard Medical School; and Ted Weiss, United States Congressman (D-N.Y.).

- Among the prominent individuals and organizations which in general support standardized testing are George H. Hanford, president of the College Board; Fred Hargadon, chairman of the College Board; William W. Turnbull, former president of the Educational Testing Service; Oluf M. Davidsen, president of the American College Testing Program; John A.D. Cooper, president of the Association of American Medical Colleges; Albert Shanker, president of the American Federation of Teachers; Gordon M. Ambach, New York Commissioner of Education; and Diane Ravitch, associate professor of history and education, Teachers College, Columbia University.
- Highlights of some of the important criticisms directed toward admissions tests and the role of the test makers, as well as responses to those criticisms, follow:

1. Standardized tests serve as gates that screen out disproportionate numbers of blacks and other minority candidates from employment and educational opportunities.

   Since each college establishes its own admissions criteria, there is no single gate in the admissions process. Standardized tests, moreover, are a means of testing *in* rather than of testing *out.*

2. Test takers are involuntary consumers who must either patronize the test makers or abandon their own educational plans.

   The institution to which the applicant applies stipulates the requirements, not the test maker. Many colleges exempt applicants from admissions tests or use other methods of evaluation.

3. Test makers are not accountable for services purchased by the public.

   Test makers are accountable to the public. Test information, including sample tests and interpretation of scores and their meanings, is widely distributed to students.

4. Since testing organizations are selling consumer products nationwide, they bear little economic resemblance to other nonprofit groups such as churches, hospitals, and schools.

   Because testing organizations are operating in the public's interest, their tax-exempt status is similar to other nonprofit organizations such as churches, hospitals, and schools.

5. Test makers' files contain vast amounts of personal, educational, and psychological information on millions of people.

   Test data and other information are provided voluntarily by test takers. Nonetheless, this information is carefully guarded to insure confidentiality; it cannot be released without the candidate's permission.

6. The costs of test development have been greatly exaggerated since it takes very little time to produce new test questions and many of these questions are reuseable.

   Test development involves detailed steps which are numerous, time-consuming, and costly, including steps to eliminate cultural and racial bias and pretesting questions before they appear on a test.

7. Because the standard error of measurement is large, test scores are at best only approximations; yet colleges make fine distinctions among candidates based on scores that are less than the standard error of measurement.

   All forms of measurement and evaluation contain some error, including grades, letters of recommendation, and personal interviews. Tests help eliminate unfair judgments arising from grades as well as the old boy/old girl network. Those who evaluate admission materials submitted by students should not make fine distinctions among the candidates' test scores.

8. Standardized tests such as the SAT do not measure aptitude, but rather how a test taker responded to a few multiple-choice questions. The results of the test are then presented as indicators of the quality of an individual's mind.

   Standardized tests such as the SAT are not designed to assess innate intelligence or unchanging abilities. Tests do not measure a person's worthiness as a human being; they are described as aptitude tests because they are not tied to a specific course of study or school curriculum.

9. Standardized tests do not measure important human qualities such as creativity, determination, stamina, idealism, wisdom and judgment.

   Standardized tests remain the best indicators of student ability regardless of a student's school curriculum, or socioeconomic background; they are not designed to measure creativity, idealism, or other important human attributes.

10. Standardized tests place blacks and other minorities at a disadvantage based on their lower than average test scores rather than their past school performance.

    While minority students may achieve lower test scores than whites, it is not the fault of the tests. Tests may reveal the inequality; they do not create it. Tests have opened doors of opportunity to minority and disadvantaged youths over the past two decades.

11. Standardized tests rank people according to their family income and discriminate between wealthy, middle, and working classes.

    Test scores do relate to family income and economic background, but many students from high income groups earn low test scores, while many students from low income groups earn high test scores.

12. Test scores add little to the prediction of college grades over the high school record alone, while other forms of assessment can predict college success nearly as well or better than test scores.

    Test scores provide a strong incremental addition to the validity of predicting college success. Moreover, test scores are not usually considered by themselves but are used in conjunction with high school grades, letters of recommendation, personal accomplishments, and personal interviews. Tests may not be perfect, but they are useful.

13. Diverse approaches are needed to assess properly a test taker's potential to perform college-level work. Because of economic and political considerations, test makers have an interest in maintaining the status quo.

    Tests do not perpetuate an unjust social system. They reduce the unfairness due to subjective assessments of personal qualities and serve to identify talent from diverse backgrounds.

Special Preparation or Coaching for College Admissions Tests

- Interest in special preparation or coaching on standardized examinations for admission to undergraduate, graduate, and professional schools nationwide continues to be strong.
- Test critics and advocates disagree on whether schools can or should attempt to improve admissions test scores. In either case, both sides emphasize that special preparation activities should not infringe on the regular high school curriculum to the exclusion of other curricular needs.
- Research findings on the effects of special preparation on college admissions examinations appear to be inconsistent because of (1) variations in the duration and quality of special preparation (2) effects of student motivation and self-selection (3) differential growth in verbal and mathematical reasoning skills over time (4) lack of perfection in the measuring instruments involved and (5) uncontrolled factors in research design. Consequently, some students might experience significant score increases while others might experience significant score declines.

- The wide variety in the types of test preparation programs currently available make it difficult to interpret the results from the research and literature on the possible effects of coaching for college admission tests. Traditionally, the terms "coaching" and "special preparation" applied only to short-term drill or practice on sample test questions for a few hours. More recently, coaching and special preparation of longer duration, up to 40 hours, has been referred to as "instruction" or "extended educational experience."
- There is general agreement that as special preparation moves from short-term drills and cramming to formal instruction of sufficient quality and duration in reading comprehension and mathematical concepts (whether through intensive independent study, additional academic courses, wide-ranging reading, or extended educational experiences such as an 8-to-10 week commercial preparatory course with homework) greater improvements can be expected in a student's admissions test score.
- Test critics maintain that a recent College Board study of the effectiveness of special preparation programs at eight private and public schools, which employed both control and treatment groups and averaged 13 hours of instruction, may have been of such short duration that meaningful improvements in test scores did not occur.
- From 1967 to 1980 the decline in SAT scores for males and females combined (42 SAT-V and 26 SAT-M points) was 68 points. The 25 test points attributable to the effects of special preparation, as estimated by the College Board correspond to approximately 37 percent of the total decline in test scores between 1967 and 1980. The 50 test questions attributable to the effects of special preparation, in a study by the Bureau of Consumer Protection of the Federal Trade Commission, correspond to approximately 74 percent of the total decline in scores between 1967 and 1980. The 100 test points attributable to the effects of special preparation, in a report by the Boston Regional Office of the FTC, exceed the total decline by 47 percent. The maximum amount of test points ever attributed to special preparation (143 test points, according to the National Education Association Study), exceeds the total decline in test scores over the 13 year period by 110 percent.
- The possibility that large numbers of students receive special preparation or coaching of an extended nature similar to "formal instruction" in commercial coaching schools or in special preparatory classes, especially in private college-preparatory high schools, calls into question the claims of test makers that standardized test scores place students from diverse geographic, socio-economic, and educational backgrounds [illegible] an "equal footing."
- There is some evidence that students who are underachievers (those who do not perform well on standardized admissions tests in light of their previous grades in school, class rank, and personal characteristics) might benefit more from special preparation than students of higher developed academic abilities.
- The possibility that coaching and other forms of special preparation for taking standardized tests for college admissions could even mildly improve the test scores of underachievers and others, under certain conditions, raises the issue of possible discrimination and the denial of equal access to educational institutions. This is especially relevant for economically and socially dis-

advantaged students who cannot afford the cost of a commercial coaching school or who do not attend a high school that offers special preparation or rigorous academic courses.

- Considering the differences in the composition of the initial standardization group of 10,654 college-bound high school seniors who were administered the SAT in April 1941 (whose mean score average was 500 and to which all subsequent groups of SAT takers have been indirectly compared), and the changing composition of today's college youth, which includes one million SAT takers and greater proportions of Blacks, Hispanics, women, and disadvantaged students than in the past, the question arises as to whether it is educationally desirable to maintain an "unchanging standard" for measuring college potential.
- When making decisions regarding the preparation of college-bound students for standardized admissions examinations, school administrators and policy makers face a basic decision: do the relative merits of attempting to improve admissions test scores through a test preparation program outweigh any possible negative effects to the institutional program for students hoping to enter colleges not relying heavily on admission test scores?
- In secondary schools planning to implement special test preparation activities there are a host of factors to consider, including whether
    1. there will be curricular effects from teaching for the tests;
    2. the activities will be elective, extracurricular, or incorporated into the regular classroom;
    3. juniors or seniors or both should participate;
    4. the activities should be of short-term or long-term duration;
    5. the program should focus on both vocabulary or mathematical skills or include other content matter;
    6. commercial review books or teacher handouts or both should be used;
    7. homework should be assigned;
    8. the program will be free or reflect charges for instructional time and materials used;
    9. there will be any political consequences for test preparation activities.
- The College Board advises secondary schools concerned with improving their students' admissions test scores to strengthen the regular curriculum and administer to groups of students the complete sample test contained in the CEEB booklet *Taking the Sat*. Some students who are not acquainted with the SAT may benefit from (1) a review of mathematical concepts, especially if they are not enrolled in a mathematics course; and (2) instruction in test-taking skills --including the format of the test, using time efficiently, knowing how to approach different types of questions, and knowing when to guess sensibly.
- The National Education Association advises that free coaching should be provided in order to assure "equity in access" for all students to the college of their choice. NEA further stresses that there are alternatives to admissions test which can better indicate a student's potential to perform college-level work. The teacher association recommends that in the admissions process college and universities should:
    1. place greater reliance on descriptive statements of students' academic achievement, strength, personal qualities and activities;
    2. place greater value on students' products in art, science, and the

performing arts;

3. assess students' work experiences and personal events;
4. encourage universities to hold discussions with students; and
5. open admissions to universities for students willing to accept a challenge.

Truth-in-Testing Legislation

- New York has passed a full test disclosure law which requires the release of test questions and answers. California has enacted a limited test disclosure law. Similar truth-in-testing legislation has been introduced, postponed, or rejected in as many as 20 states.
- Since the enactment of New York's truth-in-testing law, approximately 4.8 percent of the high school students taking the SAT in New York have requested copies of their test questions and answer sheets, compared to approximately 17 percent of those taking the LSAT.
- There is evidence that test sponsors such as the College Board and test makers such as the Educational Testing Service are trying to comply voluntarily with consumer requests for more information about tests through the adoption of a set of "Public Interest Principles" and by allowing students to buy back their SAT answer sheets and scoring keys, but not the test questions, and by returning PSAT test questions and answers to junior at no charge.
- There is some evidence that many of the legal and administrative difficulties stemming from the enactment of New York's standardized testing law have been remedied through amendments to the law. Nonetheless, persons and groups caution that hastily enacted truth-in-testing laws, with stringent provisions and diverse educational agencies entrusted to enforce the laws may seriously "tie the hands of the test makers."

Use of Test Scores in the College Admissions Process

- With the exception of those students applying to highly select colleges and universities, there is some evidence that gaining admission to college at the undergraduate level is not as difficult today as it was in the early 1970s, and may be even less difficult through the 1980s.
- High school grades are the single best predictor of college performance, with an approximate validity of .50, when both high school grades and standardized test scores are considered. There is an incremental addition of approximately .08, with the combined predictors averaging about .58.
- Colleges in general use no specific formula in making admissions decisions. Different weights, which in most cases are not precisely defined, are assigned by each college and university to various admissions criteria such as: (1) high school academic performance (2) standardized test scores (3) extracurricular activities (4) letters of recommendation from teachers and counselors (5) individual talents (6) class rank (7) personal interviews (8) student motivation, and (9) previous accomplishments.
- Statements from some college and university admissions officers suggest that more emphasis is currently placed on the overall quality and rigor of a student's high school academic program than on standardized test scores alone.
- Less than one percent of private four-year colleges and less than four percent of public four-year colleges consider test scores to be the "single most important factor" in the admissions process. Nevertheless, 54 percent of private four-year

colleges and 60 percent of public four-year colleges view admissions test scores as a "very important factor", according to a survey sponsored by the American Association of Collegiate Registrars and Admissions Officers and the College Board.

- There are some indications that in the near future assessments of an applicant's personal qualities, activities, and work experiences, which supplement test scores and grades, will play a greater role in the selection of applicants for college admission and possibly expand the admissions process.
- Provided that an applicant meets a college's minimum cutoff test score, if one exists, there is evidence that college admissions officers review all of the materials they can obtain on an individual applicant in order to assess the applicant's potential to perform college-level work.

How Test Scores Might Be Improved

- Statements by educators and officials from colleges and universities professional associations representing educators in various areas (such as reading, English, mathematics, and secondary schools) national teacher organizations, and test makers provide an array of recommendations for improving the standardized test scores of secondary school students.
- The most frequently cited recommendation by educators for improving college admission test scores is the importance of a strong secondary school curriculum, including 3-4 years of English and mathematics, broad ranging student reading, and an increased emphasis on written English.
- Other frequent recommendations for improving test scores include a serious commitment to education by parents and the community, a familiarization of students with test formats, and practice with test materials provided by the test makers.
- Some believe that special preparation, whether contained within the traditional high school curriculum or obtained through commercial coaching schools, will help students raise their test scores. But others, in particular the test makers, contend that coaching is not really necessary for students to excel on the SAT or the ACT. Instead they emphasize the importance of a strong overall curriculum experience for students.
- When considering possible responses to declining college admissions test scores and ways to increase such scores, it is important to take into account that these tests are designed to measure higher levels of reading and mathematical comprehension. School programs designed to improve college admission test scores should not be confused with programs to improve functional literacy or minimum competency.

APPENDIXES

Page

APPENDIX A -- "Application Procedures" Excerpt from the "Joint Statement on Principles of Good Practice in College Admissions and Recruitment," Developed by the American Association of Collegiate Registrars and Admissions Officers, the College Board, the National Association of College Admissions Counselors, and the National Association of Secondary School Principals. 140

APPENDIX B -- The Mathematical Association of America and the National Council of Teachers of Mathematics Position Statement on "Recommendations for the Preparation of High School Students for College Mathematics Courses" 142

APPENDIX C -- National Council of Teachers of Mathematics "Recommendations for School Mathematics of the 1980s" 144

APPENDIX D -- "National Testing Movement": Basic Policy Statement, Resolutions on Education, National Association for the Advancement of Colored People, 1979 145

APPENDIX E -- Operational Elements of "Public Interest Principles for the Design and Use of Admissions Testing Programs," Proposed by the College Board, Educational Testing Service, Graduate Management Admission Council, Graduate Record Examinations Board, and Law School Admission Council 146

APPENDIX A

"Application Procedures" Excerpted from the "Joint Statement on Principles of Good Practice in College Admissions and Recruitment," Developed by the American Association of Collegiate Registrars and Admissions Officers, the College Board, the National Association of College Admissions Counselors, and the National Association of Secondary School Principals

---

1. Colleges and universities will:

   A. Accept full responsibility for admissions decisions and for proper notification of those decisions to candidates and, when possible, to secondary schools.

   B. Receive information about candidates in confidence and respect completely, within the confines of federal and/or state laws, the confidential nature of such data.

   C. Not apply newly revised requirements to the disadvantage of a candidate whose secondary school course has been established in accordance with earlier requirements.

   D. Notify candidates as soon as possible if they are clearly inadmissible.

   E. Not deny admission to a candidate on the grounds that their institution does not have aid funds to meet the candidate's apparent financial need, except for foreign students.

   F. Not require candidates or their schools to indicate the order of candidates' college or university preferences, except under early-decision plans.

   G. Permit candidates to choose, without penalty, among offers of admission until they have heard from all colleges to which they have applied, or until the date established under the Candidates Reply Date Agreement.

   H. Maintain a waiting list of reasonable length and only for a reasonable period of time.

   I. State clearly the application procedures for transfer students by informing candidates of deadlines, documents required, courses accepted, and course equivalency.

2. Secondary schools will:

   A. Provide for colleges and universities accurate, legible, and complete transcripts for their school's candidates.

   B. Describe their school's marking system and method of determining rank in class.

   C. Describe clearly special curricular opportunities (e.g., honors, advanced placement courses, seminars, etc.).

   D. Provide accurate descriptions of the candidates' personal qualities that are relevant to the admissions process.

   E. Report any significant change in candidates' status or qualifications between the time of recommendation and graduation.

   F. Urge candidates to recognize and discharge their responsibilities in the admissions process by:

      - complying with requests for additional information in a timely manner;

      - responding to institutional deadlines on admissions and refraining from stockpiling acceptances;

      - responding to institutional deadlines on room reservations, financial aid, health records, and prescheduling where all or any of these are applicable.

G. Not, without permission of candidates, reveal the candidates' college preference.

H. Advise students not to sign any contractual agreement with an institution without examining the provisions of the contract.

I. Advise students to notify other institutions when they have accepted an admissions offer.

3. Community agencies will:

A. Exercise their responsibility to the entire educational community.

B. Discourage unnecessary multiple applications.

C. Discourage students from stockpiling offers of admission.

SOURCE: American Council on Education. "Self-Regulation Initiatives: Guidelines for Colleges and Universities." Washington, DC: ACE, October 1979, p. 2.

APPENDIX B

The Mathematical Association of America and the National Council of Teachers of Mathematics Position Statement on "Recommendations for the Preparation of High School Students for College Mathematics Courses."

The Board of Governors of the Mathematical Association of America and the Board of Directors of the National Council of Teachers of Mathematics make the following recommendations:

1. Proficiency in mathematics cannot be acquired without individual practice. We, therefore, endorse the common practice of making regular assignments to be completed outside the class. We recommend that parents encourage their children to set aside sufficient time each day to complete these assignments and that parents actively support the request of the teachers that homework be turned in. Students should be encouraged to develop good study habits in mathematics courses at all levels and should develop the ability to read mathematics.

2. Homework and drill are very important pedagogical tools used to help the student gain understanding as well as proficiency in the skills of arithmetic and algebra; but students should not be burdened with excessive or meaningless drill. We, therefore, recommend that teachers and authors of textbooks step up their search for interesting problems that provide the opportunity to apply these skills. We realize that this is a difficult task, but we believe that providing problems that reinforce manipulative skills as a by-product should have high priority, especially those that show that mathematics helps solve problems in the real world.

3. We are aware that teachers must struggle to maintain standards of performance in courses at all levels from kindergarten through college and that serious grade inflation has been observed. An apparent growing trend to reward effort or attendance rather than achievement has been making it increasingly difficult for mathematics teachers to maintain standards. We recommend that mathematics departments review evaluation procedures to insure that grades reflect student achievement. Further, we urge administrators to support teachers in this endeavor.

4. In light of 3 above, we also recognize that advancement of students without appropriate achievement has a detrimental effect on the individual student and on the entire class. We, therefore, recommend that school districts make special provisions to assist students when deficiencies are first noted. [emphasis in the original]

5. We recommend that cumulative evaluations be given throughout each course, as well as at its completion to all students. [emphasis in the original] We believe that the absence of cumulative evaluation promotes short-term learning. We strongly oppose the practice of exempting students from evaluations.

6. We recommend that computers and hand calculators be used in imaginative ways to reinforce learning and to motivate the student as proficiency in mathematics is gained. Calculators should be used to supplement rather than to supplant the study of necessary computational skills.

7. We recommend that colleges and universities administer placement examinations in mathematics prior to final registration to aid students in selecting appropriate college courses.

8. We encourage the continuation or initiation of joint meetings of college and secondary school mathematics instructors and counselors in order to improve communication concerning mathematics prerequisites for careers, preparation of students for collegiate mathematics courses, joint curriculum coordination, remedial programs in schools and colleges, an exchange of successful instructional strategies, planning of in-service programs, and other related topics.

9. Schools should frequently review their mathematics curricula to see that they meet the needs of their students in preparing them for college mathematics. School districts that have not conducted a curriculum analysis recently should do so now, primarily to identify topics in the curriculum which could be either omited or de-emphasized, if necessary, in order to provide sufficient time for the topics included in the above statement. We suggest that, for example, the following could be de-emphasized or omitted if now in the curriculum:

   a. logarithmic calculations that can better be handled by calculators or computers,

   b. extensive solving of triangles in trigonometry,

   c. proofs of superfluous or trivial theorems in geometry.

10. We recommend that algebraic concepts and skills be incorporated wherever possible into geometry and other courses beyond algebra to help students retain these concepts and skills.

SOURCE: The Mathematical Association of America and the National Council of Teachers of Mathematics. "Recommendations for the Preparation of High School Students for College Mathematics Courses." Washington, DC: MAA, n.d., p. 4.

APPENDIX C

National Council of Teachers of Mathematics "Recommendations for School Mathematics of the 1980s"

The National Council of Teachers of Mathematics recommends that--

1. problem solving be the focus of school mathematics in the 1980s;
2. basic skills in mathematics be defined to encompass more than computational facility;
3. mathematics programs take full advantage of the power of calculators and computers at all grade levels;
4. stringent standards of both effectiveness and efficiency be applied to the teaching of mathematics;
5. the success of mathematics programs and student learning be evaluated by a wider range of measures than conventional testing;
6. more mathematics study be required for all students and a flexible curriculum with a greater range of options be designed to accommodate the diverse needs of the student population;
7. mathematics teachers demand of themselves and their colleagues a high level of professionalism; and
8. public support for mathematics instruction be raised to a level commensurate with the importance of mathematical understanding to individuals and society.

SOURCE: The National Council of Teachers of Mathematics. *An Agenda for Action: Recommendations for School Mathematics of the 1980s.* Reston, VA: NCTM, 1980, p. 1. Copyright 1980 by The National Council of Teachers of Mathematics, Reston, VA.

APPENDIX D

"National Testing Movement": Basic Policy Statement, Resolutions on Education, National Association for the Advancement of Colored People, 1979

Whereas, the use of tests is escalating in our society and standardized tests in some instances are proposed as the sole criteria for nonpromotion of students; and,

Whereas, state legislation authorizing competency testing for high school students has spread to 36 states and proposals are being discussed at the Federal level to require a national test for high school graduation; and,

Whereas, increased use of testing is being imposed with changes in the curriculum, inservice teacher training or the effective use of diagnostic-prescriptive measures to assess student needs; and,

Whereas, such untrammeled use of testing instruments is another way of blaming the student victim;

Therefore, be it resolved, that the NAACP work for truth in testing legislation at the Federal level;

Be it also resolved, that the NAACP reaffirm its policy demanding a moratorium on standardized testing wherever such tests have not been corrected for cultural bias; and,

Be it further resolved, that the NAACP adamantly oppose the use of testing results in an adverse fashion and any movement geared to the use of scores on a national test as prerequisite for high school graduation.

Be it also resolved, that we direct our units to request specific information from schools re: courses and programs including goals and objectives, planned activities and substance; measures used for evaluation, expectations for student and teacher competencies and sample assignments required for success in the course or program.

Be it finally resolved, that the national office develop a positive national strategy to implement this resolution.

SOURCE: National Association for the Advancement of Colored People. "National Testing Movement." Basic Policy Statement, Resolutions on Education. Washington, DC: NAACP, 1979.

APPENDIX E

Operational Elements of "Public Interest Principles for the Design and Use of Admissions Testing Programs," Proposed by the College Board, Educational Testing Service, Graduate Management Admission Council, Graduate Record Examinations Board, and Law School Admission Council

The separately constituted and governed groups sponsoring testing programs may choose to implement these principles in different ways. This probable diversity stems from differences in the nature and purposes of the tests in the several programs and from the specifics of their structure and operation. Examples of possible approaches include the following:

1. Each prospective examinee should be able to receive a full-length sample of each test, similar to the one he or she will take, with the intended answers and with instructions for self-administration and self-scoring.
2. For tests given to a sufficient number of students annually to support the cost, at least one operational form of the test should be published periodically, in addition to the regular sample. A specific schedule of publication should be designated for each program.
3. Non-technical information about the testing program should be furnished routinely to test-takers, users, and the general public. It should include a description of what each test measures, the error of measurement, how the scores are intended to be used, and a summary of the validity of the scores for the intended uses.
4. A technical publication should provide information on the same topics in sufficient depth to permit professionals in the field to assess the evidence and the accuracy of the non-technical summary.
5. Studies of the use of the test by professionals other than those in the sponsoring or administering agency should be actively encouraged and facilitated by provision of the necessary data with safeguards for individual privacy. The results of those studies should be published in regular journals and also incorporated in the technical and non-technical publications.
6. The test sponsor should ensure that operational forms of the tests are independently reviewed before they are given. The review should include the appropriateness of the content of the test and in particular should seek to detect and remove potential racial, cultural or sex bias or other influences extrinsic to the characteristics, skills or knowledge to be measured. The review should also determine that the operational form is fairly represented by the sample test already distributed.
7. Test takers should have the right to question the accuracy of scoring, administrative procedures, specific questions in a test, or allegations of irregularities in test administrations. Current procedures to deal with this right should be reviewed and modified if necessary to ensure a fair and prompt response.

We hope communication of these principles and operational guidelines leads to greater understanding and constructive dialogue about the important issues surrounding testing. We stand ready to work with all interested groups in discussion of the policies and improvement of the procedures under which testing programs are conducted.

SOURCE: Educational Testing Service. "Public Interest Principles for the Design and Use of Admissions Testing Programs." Princeton, NJ: ETS, January 2, 1980, p. 3.

REFERENCES

The following bibliography is based on a search of published and unpublished literature. All of the entries relate to standardized testing, although not all are cited in the text.

1. "AAMC Challenges in Federal Court New York Standardized Testing Law." Washington, DC: Association of American Medical Colleges (One Dupont Circle, Suite 200, 20036), November 9, 1979. 3 pp. (news release)

2. "AAMC Takes Truth-in-Testing Law to Court," *Education Daily*, 12 (November 9, 1979), p. 5.

3. "ACT Assessment Test Scores--3rd Year in Row for Increase," *Activity*, 18 (May 1980), p. 4. (A publication of the American College Testing Program, Iowa City, IA)

4. "Action on Testing Bill Unlikely Soon," *Education Daily*, 12 (October 26, 1979), p. 6.

5. "On Admissions Testing: A Conversation with Fred Hargadon," *Educational Leadership*, 37 (May 1980), pp. 655-657.

6.0 Admissions Testing Program of the College Board. *ATP Guide for High Schools and Colleges 1979-81.* Princeton, NJ: CEEB (Box 592, 08541), 1979, 1980. 28 pp.

6.1 ________________________________. *Guide to the Admissions Testing Program 1978-79.* Princeton, NJ: College Entrance Examination Board, 1978. 30 pp.

6.2 ________________________________. *National College-Bound Seniors, 1980.* New York, NY: CEEB (888 Seventh Ave., 10019), 1980. 22 pp.

7. ________________________________. *National College-Bound Seniors, 1979.* New York, NY: CEEB, 1979 . 22 pp.

8. ________________________________. *National Report: College-Bound Seniors, 1978.* New York, NY: CEEB, 1978. 22 pp.

9. ________________________________. *National Report: College-Bound Seniors, 1977.* New York, NY: CEEB, 1977. 21 pp.

10. ________________________________. *National Report: College-Bound Seniors, 1975-76.* New York, NY: CEEB, 1975, 1976. 21 pp.

11. ________________________________. *Taking the SAT: A Guide to the Scholastic Aptitude Test and the Test of Standard Written English.* Princeton, NJ: CEEB (P.O. Box 2815), 1978. 46 pp.

12. "AFT Solidifies Pro Testing, Anti Truth-in-Testing Stance," American Federation of Teachers *News Release*, June 5, 1980. 2 pp.

13. Alderman, Donald L. and Donald E. Powers. *The Effects of Special Preparation on SAT-Verbal Scores.* Princeton, NJ: Educational Testing Service (08541), February 1979. 37 pp.

14. Ambach, Gordon M. "Statement of Gordon M. Ambach, President of the University of the State of New York and Commissioner of Education." Paper presented to the New York State Senate and Assembly Higher Education Committee, May 9, 1979. 2 pp.

15. The American College Testing Program. *ACT Class Profile Report Enrolled 1979-80 Freshmen: National Norms 10% Sample.* Iowa City, IA: ACT (P.O. Box 168, 52243), 1980. 27 pp.

16. ____________________. *ACT General Information.* Iowa City, IA: ACT, 1977. 28 pp.

17. ____________________. "ACT Statement on Coaching." Iowa City, IA: ACT, October 1979. 6 pp.

18.0 ____________________. *Using the ACT Assessment on Campus.* Iowa City, IA: ACT, 1978. 36 pp.

18.1 American Council on Education. "Self-Regulation Initiatives: Guidelines for Colleges and Universities." Washington, DC: ACE (One Dupont Circle, 20036), October 1979. 3 pp.

19. Angoff, William H. (ed.). *The College Board Admissions Testing Program: A Technical Report on Research and Development Activities Relating to the Scholastic Aptitude Test and Achievement Tests.* Princeton, NJ: College Entrance Examination Board (Box 592, 08540), 1971. 181 pp. ($5.00)

20. Applebee, Arthur N. "The SAT Score Decline Report," *Slate*, 2 (November 1977), pp. 1-3. (A publication of the National Council of Teachers of English, Urbana, IL 61801)

21. Armstrong, Jane M. *Achievement and Participation of Women in Mathematics: An Overview.* Denver, CO: Educational Commission of the States (1860 Lincoln St., Suite 300, 80295), March 1980. 35 pp.

22. "Avoid 'Junk' Courses, Admissions Officer Says," *Education U.S.A.*, 22 (January 28, 1980), p. 164.

23. "Back to Basics? No," *Education Summary*, (April 1, 1978), pp. 1-2.

24. "Battle Brewing Over Testing Bills," *Education Daily*, 12 (September 11, 1979), p. 2.

25. Belliveau, Jeannette. "Test Scores Drop Worst in 5 Years," *Montgomery (MD) Journal*, 7 (September 26, 1979), pp. A1, A11.

26. Berkowe, Marc. "Experimental Course Boosts College Bound Scores Dramtically." Washington, DC: The Heights Study Center (4300 Garrison St., NW, 20016), February 8, 1978. 3 pp.

27. Black, John A. "When Is a Study Not a Study? An Examination of Guidelines for Improving SAT Scores," *Phi Delta Kappan*, 60 (March 1979), pp. 487-489.

28. Blum, Gerrie. "New York PTA Legislative Memo: Truth-in-Testing." Paper presented to members of the New York State Senate and Assembly, June 6, 1979. 2 pp.

29. Blumenthal, Ralph. "Test-Makers Face New Tests," *New York Times*, (April 20, 1980), pp. 1, 25.

30. "Boyer Challenges Educators to Scrutinize Testing," *Education Daily*, 12 (October 30, 1979), p. 3.

31. Brockett, Diane. "Nation's Class of '79 Gets Lowest Scores on Scholastic Aptitude Test," *Washington Star*, (September 9, 1979), p. A-3.

32. Brocklehurst, Nancy. "Women and Math: Overcoming Avoidance," *College Board Review*, 111 (Spring 1979), pp. 22-24.

33. Brown, Rexford. "Searching for the Truth about 'Truth-in-Testing,'" *Compact*, 13 (Winter 1980), pp. 7-11.

34. Brown, Rexford and Merle Steven McClung. *Searching for the Truth about 'Truth in Testing' Legislation: A Background Report.* Denver, CO: Education Commission of the States (1860 Lincoln St., Suite 300, 80295), 1980. 89 pp. ($6.50)

35. Brownstein, Ronald and Allan Nairn. "Are Truth-in-Testing Laws a Fraud? No!" *Phi Delta Kappan*, 61 (November 1979), pp. 189-191.

36. Carey, Hugh L. "Memorandum Filed with Senate Bill 5200-A, Entitled: 'An Act to Amend the Education Law, in Relation to Standardized Testing.'" Albany, NY: State of New York Executive Chamber, July 13, 1979. 2 pp.

37. ____________. "Memorandum Filed with Senate Bill Number 9142-A and Assembly Bill 8479-A." Albany, NY: State of New York Executive Chamber, July 3, 1980. 2 pp.

38. "CB Bows to Pressure, Will Release PSAT Scores," *Education U.S.A.*, 22 (April 14, 1980), p. 250.

39. "Coaching Daze: The FTC vs. Kaplan," *Time*, 113 (June 11, 1979), p. 57.

40. "Coaching, Motivation Can Help Raise SAT Scores, ETS Says," *Education Daily*, 13 (May 30, 1980), pp. 1-2.

41. "College Board Announces Fewer SAT Dates, Higher Test Fees," *Education Daily*, 12 (October 10, 1979), pp. 1-2.

42. "College Board Cuts SAT's in New York, Citing State's New Law on Testing," *Chronicle of Higher Education*, 19 (October 15, 1979), p. 2.

43. "College Board: Downslide Halts on Verbal Scores," *Guidepost*, 21 (October 5, 1978), p. 12. (A publication of the American Personnel and Guidance Association, Washington, DC)

44. "College Board Releases Early Report of College Admissions Study," *News from the College Board*, (November 1, 1979), 11 pp.

45. "College Board Ups New York SAT Charge, Adds Disclosure Fee," *Education Daily*, 13 (January 2, 1980), p. 1.

46. College Entrance Examination Board. "Backgrounder on Legislation Aimed at Government Regulation of Admissions Testing." New York, NY: CEEB (888 Seventh Ave., 10019), October 1979. 4 pp.

47. ______________________________. *College Board Annual Report, 1978-79 Cooperating to Increase Educational Opportunities.* New York, NY: CEEB, 1979. 32 pp.

48. ______________________________. *The College Board Today: A Guide to Its Programs and Services 1978.* New York, NY: CEEB, 1978. 17 pp.

49. ______________________________. "A Comparative Summary: New York and National Testing Legislation." New York, NY: CEEB, October 1979. 2 pp.

50. ______________________________. "The Effect of Special Preparation Programs on Score Results of the Scholastic Aptitude Test," *Research and Development Update.* New York, NY: CEEB, January 1979. 4 pp.

51. ______________________________. *News from the College Board*, October 10, 1979. 5 pp.

52. ______________________________. *News from the College Board*, November 19, 1979. 1 p.

53. ______________________________. "Statement of the College Board in Opposition to Senate Bill 5200/Assembly Bill 7668." New York, NY: CEEB, May 1, 1979. 2 pp.

54. ______________________________. "Statements from Individuals and Organizations Who Have Expressed Opposition and Concerns about Testing Legislation." New York, NY: CEEB, October 1979. 5 pp.

54.1 ________________________________. *Undergraduate Admissions: The Realities of Institutional Policies, Practices, and Procedures.* A report on a survey conducted by the American Association of Collegiate Registrars and Admissions Officers and the College Board. New York, NY: CEEB, 1980. 72 pp.

55. ________________________________ and National Merit Scholarship Corporation. *About Your 1979 PSAT/NMSQT Scores.* Princeton, NJ: CEEB (P.O. Box 2815, 08541), 1979 15 pp.

56. ________________________________. *1978 PSAT/NMSQT Interpretive Manual for Counselors and Administrators.* Princeton, NJ: CEEB, 1978. 22 pp.

57. "Colleges Seek New Criteria for Admissions," *ETS Developments*, 26 (1980), p. 11.

58. Committee on Post Office and Civil Service. *Professional and Administrative Career Examination.* Washington, DC: Hearing before the Subcommittee on the Civil Service of the Committee on Post Office and Civil Service, House of Representatives, Ninety-Sixth Congress, May 15, 1979. 203 pp.

59. Complaint Issued by Sullivan & Cromwell, Attorneys for College Entrance Examination Board in *CEEB* v. *Abrams*, Supreme Court of the State of New York, County of Albany, November 19, 1979. 24 pp.

60. "Could This Be Why Test Scores Have Dropped?" *American School Board Journal*, 164 (December 1977), pp. 13-14.

61. "Cram Courses Discounted as College Entrance Aid," *Washington Post*, (May 30, 1980), p. A10.

62. Crenshaw, Albert B. "College Entry Tests Faulted on Reliability, Fairness," *Washington Post*, (January 15, 1980), p. A5.

63. Davidsen, Oluf M. "ACT's Viewpoint on Testing Legislation," *Activity*, 18 (May 1980), pp. 5-8. (A publication of the American College Testing Program, Iowa City, IA)

64. Educational Research Service. *Declines in Standardized Test Scores Among Secondary School Students.* Arlington, VA: ERS (1800 N. Kent St., Suite 1020, 22209), 1976. 35 pp. ($5.00)

65. Educational Testing Service. *Accountability, Fairness, and Quality in Testing.* Princeton, NJ: ETS Information Division (08541), January 1980. 7 pp.

66. ________________________. "Coaching." Princeton, NJ: ETS, May 29, 1980. 2 pp.

67. ________________________. "The Equating of Tests." Princeton, NJ: ETS Information Division, December 11, 1979. 2 pp.

68. ________________________. *ETS 1979 Annual Report.* Princeton, NJ: ETS, 1980. 24 pp.

69. ________________________. *ETS 1978 Annual Report.* Princeton, NJ: ETS, 1978. 24 pp.

70. ________________________. "Fact and Fiction." Princeton, NJ: ETS, January 12, 1980. 3 pp.

71. ________________________. "Legislation on Standardized Testing." Princeton, NJ: ETS Information Division (08541), March 5, 1980. 5 pp.

72. ________________________. "Public Interest Principles." Princeton, NJ: ETS, January 2, 1980. 2 pp.

73. ________________________. "Public Interest Principles for the Design and Use of Admissions Testing Programs." Princeton, NJ: ETS, January 2, 1980. 3 pp.

74. ________________________. "Statement on the Federal Trade Commission Staff Report on Coaching Schools." Princeton, NJ: ETS, May 30, 1979. 3 pp.

75. __________________________. "Test Bias and Group Score Differences." Princeton, NJ: ETS, January 4, 1980. 3 pp.

76. __________________________. *Test Scores and Family Income: A Response to Charges in the Nader/Nairn Report on ETS.* Princeton, NJ: ETS, 1980. 11 pp.

77. __________________________. *Test Use and Validity: A Response to Charges in the Nader/Nairn Report on ETS.* Princeton, NJ: ETS, 1980. 27 pp.

78. "Educators, Parents Say Marijuana Disrupts Schools," *Education Daily*, 13 (January 18, 1980), pp. 3-4.

79. "Eleven-Week Course for Disadvantaged Students," *The Ranger*, (April 12, 1979), p. 4.

80. "ETS, College Board Attack Nader," *Chronicle of Higher Education*, 20 (March 17, 1980), p. 10.

81. "ETS Critical of Climate of Testing Debate," *Education Daily*, 13 (July 15, 1980), p. 4.

82. "ETS Criticizes Nader Report on Testing," *Education Daily*, 13 (March 12, 1980), p. 6.

83. "ETS Raps Nader Report for 'Faulty Statistics' and Misrepresentation," *Phi Delta Kappan*, 61 (May, 1980), p. 655.

84. "Experts Link Low Test Scores to Back-to-Basics," *Education Daily*, 12 (October 24, 1979), p. 2.

85. Fallows, James. "The Tests and the 'Brightest': How Fair Are the College Boards?" *The Atlantic*, (February 1980), pp. 37-48.

86. Farr, Roger. "Testimony of Roger Farr before Senator Thomas F. Eagleton's Hearing on the Teaching and Learning of Basic Academic Skills in Schools for the Senate Subcommittee on Education, Arts, and Humanities." Paper presented to the Senate Subcommittee on Education, Arts, and Humanities, April 9, 1979. 7 pp.

87. __________ and Jill Edwards Olshavsky. "Is Minimum Competency Testing the Appropriate Solution to the SAT Decline?" *Phi Delta Kappan*, 61 (April 1980), pp. 528-530.

88. "Federal Investigators Are Testing S.A.T.," *Executive Educator*, 1 (February 1979), p. 4.

89. Federal Trade Commission. *Effects of Coaching on Standardized Admission Examinations: Revised Statistical Analyses of Data Gathered by Boston Regional Office of the Federal Trade Commission.* Washington, DC: FTC Bureau of Consumer Protection, March 1979. 36 pp.

90. __________________________. *Staff Memorandum of the Boston Regional Office of the Federal Trade Commission: The Effects of Coaching on Standardized Admissions Examinations.* Boston, MA: Staff Memorandum of the Boston Regional Office of the FTC, September 11, 1978. 174 pp.

91. "Federal Trade Commission Reaches No Final Conclusion on Commercial Coaching." *The Examiner*, 8 (June 7, 1979), pp. 1-2.

92. "Federal Truth-in-Testing Bill Introduced Today," *Education Daily*, 12 (July 24, 1979), p. 4.

93. Feinberg, Lawrence. "How GU [Georgetown University] Selects Its Students," *Washington Post*, (April 18, 1980), pp. B1, B5.

94.0 __________________ and Dennis Collins. "Area College Entrance Test Scores Fall," *Washington Post*, (October 9, 1979), p. B2.

94.1 "Fewer Students than Expected Seek SAT Answers," *Education Daily*, 13 (September 23, 1980), pp. 1-2.

95. Fields, Cheryl M. "Coaching Courses May Help Some Students Increase SAT Scores, Study Concludes," *Chronicle of Higher Education*, 18 (June 18, 1979), pp. 3-4.

96. ________________. "College Board to Let Students Verify SAT Scores," *Chronicle of Higher Education*, 20 (April 14, 1980), p. 8.

97. ________________. "Court Delays Enforcement of N.Y. Testing Law," *Chronicle of Higher Education*, 19 (January 28, 1980), p. 2.

98. ________________. "Proposed 'Truth-in-Testing' Legislation Attacked and Defended on Capitol Hill," *Chronicle of Higher Education*, 18 (August 6, 1979), p. 11.

99. ________________. "SAT Score Improvements Laid to 'Motivation,'" *Chronicle of Higher Education*, 20 (June 9, 1980), p. 4.

100. ________________. "SAT Scores Don't Measure Scholastic Aptitude, Can Be Influenced by 'Coaching,' Report Says," *Chronicle of Higher Education*, 20 (May 27, 1980), pp. 1, 5.

101. ________________. "Social Change, School Standards, TV Viewing Among Causes Cited by Panel," *Chronicle of Higher Education*, 15 (September 6, 1977), pp. 1, 13.

102. ________________. "Suit Challenges Legality of 'Truth-in-Testing' Law," *Chronicle of Higher Education*, 19 (November 13, 1979), p. 19.

103. ________________. "Testers Agree to Make More Information Public," *Chronicle of Higher Education*, 19 (January 7, 1980), p. 5.

104. ________________ and Robert L. Jacobson. "Nader Accuses ETS of 'Fraud,' Plans Drive for Testing Reform," *Chronicle of Higher Education*, 19 (January 21, 1980), pp. 5-6.

105. Fisher, James L. "Of Testing, Truth and Mr. Nader," *Washington Star*, (February 18, 1980), p. A-15.

106. Fiske, Edward B. "College Board Test Guidelines Are Revised," *New York Times*, (February 2, 1977), p. B5.

107. ________________. "Dip in S.A.T. Score Laid to Student Caliber," *New York Times*, (August 24, 1977), p. 61.

108. ________________. "5 Exam Groups Drop Opposition to Law on Tests," *New York Times* (December 30, 1979), pp. 1, 11.

109. ________________. "Resolve the Testing Dilemma, Go to the Head of the Class," *New York Times* (October 14, 1979), p. 6E.

110. ________________. "Student Questions and Answers on 'Truth-in-Testing' Provisions," *New York Times*, (October 10, 1979), pp. A1, B3.

111. ________________. "'Truth-in-Testing' Is Still the Goal," *New York Times*, (May 11, 1980), p. 22E.

112. "FTC Releases Staff Study on Effects of Coaching on Scholastic Aptitude Test Scores," *FTC News*, (May 29, 1979), 2 pp.

113. "FTC Says Coaching on SATs May Help Some Raise Scores," American Council on Education, 28 (June 1, 1979), pp. 1-2.

114. "FTC Study Says Coaching May Improve SAT Scores," *Education Daily*, 12 (May 30, 1979), pp. 1-2.

115. Gallagher, Elizabeth. "Preparing for the S.A.T. Applications of a Research Project," *Clearing House*, 53 (January 1980), pp. 230-232.

116. Gray, H. Dennis. "The Debate on Standardized Testing," *Council for Basic Education Bulletin,* 23 (May 1979), pp. 9-12.

117. _______________. "Ralph Nader and the Reign of ETS," *Basic Education,* 24 (March 1980), pp. 7-9.

118. _______________. "Truth-in-Testing," *Basic Education,* 24 (December 1979), pp. 11-13.

119. "Guide for Raising SAT Scores Questioned: NASSP Defends Its Report," *ASCO News Exchange,* 21 (February 1979), pp. 1, 8.

120. "Guidelines for Good Education," *Council for Basic Education Bulletin,* 22 (May 1978), pp. 1-4.

121. Hanford, George H. "Memo to College Board Members, Committees and Councils." New York, NY: College Entrance Examination Board (888 Seventh Ave., 10019), February 1980. 6 pp.

122. _________________. "Memo to College Board Members, Committees and Councils." New York, NY: CEEB, June 1980. 4 pp.

123. _________________. "Testing the Tests." Remarks presented at the annual convention of the Middle States Association of Colleges and Schools, December 7, 1979. 10 pp.

124. Helming, Dennis M. *Preliminary Statistical Report on the Heights Study Center's SAT Prep Course.* Washington, DC: The Heights Study Center (4300 Garrison St., N.W., 20016), November 1979. 10 pp.

125. "Help! Teacher Can't Teach!" *Time,* 115 (June 16, 1980), pp. 54-60, 63.

126. Herndon, Terry. "The Standardized Testing Coverup." Statement presented at the annual meeting of the National Education Association, July 4, 1979. 5 pp.

127. "Higher Education," *Education Summary,* (April 15, 1980), p. 8.

128. Hildebrand, John. "Debate Heats Up Over Cramming for College Boards," *Chicago Sun-Times,* (April 27, 1979), p. 46.

129. Hill, Shirley A. "National Council of Teachers of Mathematics Recommendations for School Mathematics of the 1980s." Statement by the President of the National Council of Teachers of Mathematics, n.d. 3 pp.

130. Hobart, Thomas Y., Jr. "Testimony of Thomas Y. Hobart, Jr., President, New York State United Teachers to the New York State Assembly Standing Committee on Higher Education." Paper presented to the New York State Assembly Standing Committee on Higher Education, May 9, 1979. 5 pp.

131. Hooks, Benjamin L. "NAACP Commends Nader for Report on Educational Testing Service." Statement of Benjamin L. Hooks, National Association for the Advancement of Colored People, January 14, 1980. 1 p.

132. *Initial Impact Report of the Admission Testing Legislation.* Albany, NY: The University of the State of New York, the State Education Department (12234), 1979. 19 pp.

133. Irby, Alice J. "Statement of Alice J. Irby before the Constitutional and Administrative Law Committee, Maryland House of Delegates, House Bills 1406 and 1550." Paper presented to the Maryland House of Delegates, March 6, 1980. 12 pp.

134.0 "Is Truth in Testing Constitutional? Med Schools Say No; File Federal Suit," *Phi Delta Kappan,* 61 (January 1980), pp. 366-367.

134.1 Jackson, Rex. "The Scholastic Aptitude Test: A Response to Slack and Porter's 'Critical Appraisal,'" *Harvard Educational Review,* 50 (August 1980), pp. 382-391.

135. Jacobson, Robert L. "Blacks Lag in SAT Scores," *Chronicle of Higher Education*, 19 (January 7, 1980), p. 5.

136. ________________. "Can 'Coaching' Help Students Score Higher on the S.A.T.'s?" *Chronicle of Higher Education*, 18 (April 16, 1979), pp. 1, 9.

137. ________________. "Few Students Using N.Y. Law to Obtain Test Answers," *Chronicle of Higher Education*, 20 (July 28, 1980), p. 3.

138. ________________. "New York's 'Truth-in-Testing' Law Called Threat to Quality of the SAT," *Chronicle of Higher Education*, 19 (November 13, 1979), pp. 1, 19.

139. ________________. "Standardized Testing and Cultural Bias," *Chronicle of Higher Education*, 14 (July 25, 1977), pp. 3-4.

140. ________________. "'Truth-in-Testing': More Harm Than Good?" *Chronicle of Higher Education*, 19 (November 5, 1979), p. 10.

141. Jencks, Christopher. "Why Students Aren't Learning," *Center Magazine*, (July-August 1979), pp. 12-14.

142. ________________. "The Wrong Answer for Schools Is: (b) Back to Basics," *Washington Post*, (February 19, 1978), pp. C1, C4-C5.

143. Johnston, Lloyd D., Jerald G. Bachman and Patrick M. O'Malley. *Drugs and the Nation's High School Students: Five Year National Trends.* Rockville, MD: National Institute on Drug Abuse (5600 Fishers Lane, 20857), 1979. 80 pp.

144. Judy, Stephen. "On Second Thought: Reviewing the SAT Decline," *English Journal*, 66 (November 1977), pp. 5-7.

145. Kennedy, Shawn G. "Albany Bill Challenging Secrecy in Testing," *New York Times*, (May 27, 1979), p. 35.

146. Kintisch, Lenore S. "Classroom Techniques for Improving Scholastic Aptitude Test Scores," *Journal of Reading*, 22 (February 1979), pp. 416-419.

147. Kokus, Nancy. *Revision of Grading Scale--Report to School Board.* Fairfax, VA: Fairfax County Public Schools (10700 Page Ave., 22030), January 1979. 23 pp.

148. "Lack of Academic Courses Caused SAT Score Decline: NAASP," *Education Summary*, (May 1, 1978), p. 7.

149. LaValle, Kenneth P. "Memorandum in Support of an Act to Amend the Education Law, in Relation to Standardized Testing, S-5200A." Albany, NY: The Senate (State of New York, 12247), n.d. 2 pp.

150. ________________. "Memorandum in Support of an Act to Amend the Education Law, in Relation to Standardized Testing, S-9142." Albany, NY: The Senate (State of New York, 12247), n.d. 2 pp.

151. ________________. Press release in response to the withdrawal of the Association of American Medical Colleges and the American Dental Association from the State of New York Albany, NY: Legislative Office Building (Room 606, 12247), July 19, 1979. 2 pp.

152. ________________. "Remarks to Joint Committee Hearing on Truth-in-Testing Legislation." Albany, NY: The Senate (State of New York, 12247), May 9, 1979. 2 pp.

153. ________________. "Truth-in-Testing," *Today's Education*, 69 (February-March), 1980, pp. 59-60, 62, 64.

154. Letter from Raj K. Chopra, Superintendent of Schools, Council Bluffs, IA, June 27, 1980.

155. Letter from Roger Farr, President of the International Reading Association and Professor of Education at Indiana University, Bloomington, IN, June 27, 1980.

156. Letter from Willard H. McGuire, President, National Education Association, Washington, DC, July 17, 1980.

157. Letter from Vito Perrone, Dean of the Center for Teaching and Learning, the University of North Dakota, Grand Forks, ND, July 15, 1980.

158. Letter from Alan C. Purves, President of the National Council of Teachers of English and Principal of University High School at the University of Illinois, Urbana, IL, July 15, 1980.

159. Letter from Diane Ravitch, Associate Professor of History and Education, Teachers College, Columbia University, New York, NY, July 16, 1980.

160. Letter from Philip R. Rever, Director, Washington DC Office of the American College Testing Program, July 7, 1980.

161. Letter from Sarah Theurkauf, Public Affairs Division, College Entrance Examination Board, New York, NY, January 25, 1980.

162. Letter from Scott D. Thomson, Executive Director, The National Association of Secondary School Principals, Reston, VA, August 25, 1980.

163. Letter from Sheila Tobias, author of *Overcoming Math Anxiety*, Arlington, VA, August 21, 1980.

164. Letter from James G. Ward, Director of Research, American Federation of Teachers, Washington, DC, August 12, 1980.

165. Levine, Arthur E. "The Standardized Testing Coverup." Statement presented at the annual convention of the National Education Association, July 4, 1979. 2 pp.

166. Levy, Steven. "E.T.S. and the 'Coaching' Cover-Up," *New Jersey Monthly*, (March 1979), pp.

167. Maeroff, Gene I. "Agency Finds Coaching Can Raise S.A.T. Scores," *New York Times*, (May 30, 1979), p. A10.

168. ______________. "Pick One: S.A.T. Coaching Is (Is Not) Fair to Students," *New York Times*, (May 16, 1979), p. E11.

169. "Marijuana Could Hurt Schoolchildren, Senate Told," *Education Daily*, 13 (January 17, 1980), pp. 1, 4.

170. "Marijuana Use Among Students Levels Off, Study Shows," *Education Daily*, 13 (January 17, 1980), pp. 1-2.

171. Masters, Kim. "Critics Rail as E.T.S. Power and Influence Proliferate," *Executive Educator*, 1 (May 1979), pp. 22-24.

172. The Mathematical Association of America and the National Council of Teachers of Mathematics. "Recommendations for the Preparation of High School Students for College Mathematics Courses." Washington, DC: MAA (1529 Eighteenth St., NW, 20036), n.d. 6 pp.

173. Mathews, Jessica Tuchman. "A Measure of SATs," *Washington Post*, (February 7, 1980), p. A-19.

174. Maxwell, John C. "The NASSP Guidelines for Improving SAT Scores: A Dramatic Abuse of a Testing Situation," *Slate*, 3 (September 1978), pp. 1-3. (A publication of the National Council of Teachers of English, Urbana, IL)

175. Meislen, Richard J. "New York to Open Tests for College to Public Scrutiny," *New York Times*, (July 15, 1979), Section 1, p. 1.

176. "Memorandum in Opposition." Paper submitted by Legislative Counsel on behalf of Educational Testing Service. Albany, NY: Hinman, Straub, Pigors, and Manning (90 State St., 12207), May 14, 1979. 4 pp.

177. Messick, Samuel and Others. "The Effectiveness of Coaching for the SAT: Review and Reanalysis of Research from the Fifties to the FTC: Summary of Issues and Results." Washington, DC: Educational Testing Service (1800 Massachusetts Ave., NW), May 1980. 9 pp.

178. Micklos, John. "International Reading Association President Testifies on Reading Scores," *International Reading Association News*, (June 11, 1979). 4 pp.

179. Middleton, Lorenzo. "Verbal S.A.T. Scores Leveled Off This Year," *Chronicle of Higher Education*, 17 (September 18, 1978), p. 2.

180. Molotsky, Irvin. "Truth-in-Testing Law Opposed in Congress by Makers of Exams," *New York Times*, (October 28, 1979), p. 28.

181. Myers, Catherine. "N.Y. State Enacts 'Truth-in-Testing' Laws; 2 Admissions Groups to Halt Examinations," *Chronicle of Higher Education*, 18 (July 23, 1979), pp. 1, 6.

182. _______________. "Verbal, Math SAT Scores Decline to All-Time Low," *Chronicle of Higher Education*, 19 (September 17, 1979), p. 6.

183. "Nader, ETS Clash Over Standardized Testing," *Education Daily*, 13 (January 15, 1980), p. 3.

184. Nader, Ralph. "The Pervasive Power of ETS." Excerpted from an appearance by Ralph Nader on *The Phil Donahue Show*, May 2, 1979. Cincinnati, OH: Phil Donahue Transcripts (P.O. Box 2111, 45202). 2 pp. ($2.50)

185. "Nader Report Knocks ETS Tests; Links Scores to Family Income," *Phi Delta Kappan* (March 1980), p. 504.

186. Nairn, Allan and Others. "Class in the Guise of Merit," *Educational Leadership*, 37 (May 1980), pp. 651-653.

187. ______________________. "The Reign of ETS," *Today's Education*, 69 (April-May 1980), pp. 58G-64GE.

188. ______________________. *The Reign of ETS: The Corporation That Makes Up Minds.* Washington, DC: The Learning Research Project (P.O. Box 19312, 20036), 1980. 554 pp. ($30.00)

189. "NASSP's Guidelines for Improving SAT Scores Spark a Dispute," *Education Summary*, (August 15, 1978), p. 5.

190.0 National Academy of Education. *Improving Educational Achievement: Report of the National Academy of Education to the Assistant Secretary for Education.* Washington, DC: NAE (11 Dupont Circle, Suite 130, 20036), 1978. 18 pp.

190.1 National Association for the Advancement of Colored People. "National Testing Movement." Basic Policy Statement, Resolutions on Education. Washington, DC: NAACP, 1979.

191. National Council of Teachers of Mathematics. *An Agenda for Action: Recommendations for School Mathematics of the 1980s.* Reston, VA: NCTM (1906 Association Dr., 22091), 1980. 29 pp.

192. National Education Association. "The Coachability Question." Washington, DC: NEA (1201 16th St., NW, 20036), n.d. 1 p.

193. National Education Association. *On Further Examination of 'On Further Examination.'* Washington, DC: NEA, 1977. 28 pp.

194. ______________________________. *Measurement and Testing: An NEA Perspective.* Washington, DC: NEA, 1980. 97 pp.

195. ______________________________. "NEA Calls for Congressional Investigation of 'Secretive' Multimillion Dollar Testing Industry." Statement presented at the annual meeting of the National Education Association, July 4, 1979. 3 pp.

196. ______________________________. "Truth-in-Testing: NEA's Position." Washington, DC: NEA Government Relations, October 1979. 1 p.

197. "NEA Calls for Test Coaching Study," *Education Daily*, 12 (October 17, 1979), pp. 3-4.

198. "NEA Demands Release of Standardized Testing Data," *NEA Now*, (April 2, 1979), p. 1.

199. "NEA Finds Coaching Raises SAT Scores," *Education Daily*, 13 (July 10, 1980), p. 5.

200. "NEA Opposes Regulation of Pre-College Tests," *Education Daily*, 12 (October 11, 1979), p. 3.

201. "NEA's Second Kit: Standardized Testing," *NEA Now* (January 7, 1980), p. 1.

202. "New Report Says Coaching Boosts SAT Scores," *Education Daily*, 13 (May 16, 1980), pp. 1-2.

203. "New Services for Students Announced by the College Board," *News from the College Board*, (April 8, 1980). 4 pp.

204. "New Students, Lower Standards Caused Test Score Decline," *Education U.S.A.*, 19 (August 29, 1977), pp. 377, 383.

205. "New York Assembly Passes Bill Mandating Sunday Exams," *Education Daily*, 13 (February 7, 1980), p. 5.

206. "New York Law Forces Withdrawal of New MCAT from State." Washington, DC: Association of American Medical Colleges (One Dupont Circle, Suite 200, 20036), July 17, 1979. 2 pp. (news release)

207. "New York Opens College Admissions Tests to Public Scrutiny," *Education Daily*, 12 (July 17, 1979), p. 2.

208. "New York Passes Truth-in-Testing Law; Test Sponsors Threaten to Withdraw," *Phi Delta Kappan*, 61 (September 1979), p. 77.

209. New York Public Interest Research Group, Inc. "Statement of Factual Correction and Rebuttal to Educational Testing Service Memorandum in Opposition." Albany, NY: NYPIRG (1 Columbia Place, 12207), May 18, 1979. 8 pp.

210. "New York Testing Law Temporarily Blocked," *Education Daily*, 13 (January 23, 1980), p. 3.

211. "New York 'Truth-in-Testing' Bill Makes Strong Showing," *Education Daily*, 12 (May 25, 1979), p. 4.

212. "NIE to Study Testing's Impact on College Admissions," *Education Daily*, 13 (March 14, 1980), p. 4.

213. "N.Y. Law Requires Publishing Answers to Standard Tests," *Washington Post*, (July 16, 1979), p. A20.

214. "N.Y. Legislator Warns of Test Industry's 'Tough' Lobbying," *Education Daily*, 12 (August 3, 1979), p. 2.

215. Omang, Joanne. "College Truth-in-Testing Bill Flunks Out," *Washington Post*, (October 24, 1979), p. A4.

216. ____________. "Making the Grade: Standardized Tests Are Under Attack," *Washington Post*, (December 26, 1979), pp. A10-A11.

217. ____________. "SAT Scores, Following Trend, Defy Predictions, Drop Again," *Washington Post*, (September 9, 1979), p. A12.

218. "An Overview of Findings from the College Board-AACRAO Survey of Undergraduate Admissions, Policies, Practices and Procedures." Prepared for the annual meeting of the College Board, October 30, 1979. 18 pp.

219. Perkins, James A. and Others. *Strength Through Wisdom: A Critique of U.S. Capability.* A report to the President from the President's Commission on Foreign Language and International Studies, November 1979. 39 pp.

220. Perrone, Vito. "How Not to Improve Educational Achievement," *Education Digest*, 45 (September 1979), pp. 20-23.

221. ____________. "Testimony Prepared for the Senate and Assembly Higher Education Committees Regarding the Examination and Assessment of Standardized Testing in Light of Legislation Currently Before the Senate and Assembly of the State of New York." Paper presented to the New York State Senate and Assembly Higher Education Committees, n.d. 4 pp.

222. Pike, Lewis W. Letter to Senator Kenneth P. LaValle, Chairman, Senate Higher Education Committee, Albany, NY, June 12, 1979. 2 pp.

223. ____________. *Short-Term Instruction, Testwiseness, and the Scholastic Aptitude Test: A Literature Review with Research Recommendations.* Princeton, NJ: College Entrance Examination Board (Box 2815, 08541), 1979. 44 pp.

224. ____________. "Testimony of Dr. Lewis W. Pike Before the New York State Senate and Assembly Higher Education Committees." Paper presented to the New York State Senate and Assembly Higher Education Committees, May 9, 1979. 5 pp.

225. ____________ and Franklin R. Evans. *The Effects of Special Instruction for Three Kinds of Mathematics Aptitude Items.* Princeton, NJ: Educational Testing Service (08541), May 1972. 86 pp.

226. Pooler, Rosemary. "Memorandum to Members of the Legislature." Paper presented by Rosemary Pooler to the New York Legislature, June 11, 1979. 2 pp.

227. Powers, Donald E. and Donald L. Alderman. *The Use, Acceptance, and Impact of Taking the SAT-A Test Familiarization Booklet.* Princeton, NJ: Educational Testing Service (08541), February 1979. 81 pp.

228. "Public Interest Principles," *ETS Developments*, 26 (1980), pp. 1, 9.

229. "Public Must Confront Urgent Problems, Mathematics Teachers Say." Reston, VA: National Council of Teachers of Mathematics, April 16, 1980. 4 pp.

230. Raspberry, William. "Truth-in-Testing," *Washington Post*, (September 10, 1979), p. A27.

231. Ravitch, Diane. "Are Truth-in-Testing Laws a Fraud? Yes!" *Phi Delta Kappan*, 61 (November 1979), pp. 189-190.

232. "Regents Adopt Regulations for Admissions Testing Legislation," *New York State Education Department News* (November 15, 1979). 2 pp.

233. "Regents Call for Modified Testing Law," *Education Daily*, 12 (December 20, 1979), p. 5.

234. "Regents Propose Review of Admissions Testing Legislation," *New York State Education Department News* (December 14, 1979). 2 pp.

235. "The Righting of Writing," *Time*, 115 (May 19, 1980), pp. 88-91.

236. Robertson, Dario F. "Examining the Examiners: The Trend Toward Truth in Testing," *Journal of Law and Education*, 9 (April 1980), pp. 167-199.

237. Rubin, Nancy. "Test-Coaching: Is It Worth It?" *New York Times*, (September 9, 1979), Section 12, pp. 1, 20.

238. Ryan, Robert. "College Entrance Exams Flunk Nader Test," *Miami Herald*, (January 15, 1980), p. 4A.

239. Ryor, John. "Declining SAT Scores," *Today's Education*, 66 (November-December 1977), pp. 6, 8.

240. "The S.A.T. Score Decline," *Council for Basic Education Bulletin*, 22 (October 1977), pp. 1-4.

241. "SAT Score Decline Raises Questions About Competency-Based Education," *Education Daily*, 12 (September 12, 1979), p. 2.

242. "SAT Score Drop Laid to A-Bomb," *Washington Star*, (August 29, 1979), p. A10.

243. "SAT Scores and Cramming," *Council for Basic Education Bulletin*, 23 (March 1979), pp. 8-9.

244. "SATs Drop Again," *Education Daily*, 12 (September 10, 1979), p. 1.

245. Savage, David G. "The Long Decline in SAT Scores," *Educational Leadership*, 35 (January 1978), pp. 290-293.

246. "Scholastic Aptitude Test Scores Drop as 10-Year Trend Goes On," *New York Times*, (September 9, 1979), p. 36.

247. "Schools Accept FTC Conclusions, Testing Official Says," *Education Daily*, 12 (November 7, 1979), p. 2.

248. "Scoring the Admissions Tests," *Washington Post*, (August 27, 1979), p. A26.

249. "Secondary School Grade Inflation," *NASSP Bulletin*, 63 (April 1979), p. 30.

250. "SED Studies Impact of Admissions Testing Legislation," *New York State Education Department News*, (October 18, 1979). 2 pp.

251. Shah, Diane K. and Others. "SAT Cramming: Does It Work?" *Newsweek*, (June 11, 1979), p. 113.

252. Shane, Harold G. "The Academic Score Decline: Are Facts the Enemy of Truth?" *Phi Delta Kappan*, 59 (October 1977), pp. 83-86, 145-146.

253. Shanker, Albert. "The Truth About 'Truth-in-Testing' Laws," *New York Times*, (April 20, 1980), p. E7.

254. Shapiro, Margaret. "P. G., Montgomery Scores Are Down on College Boards," *Washington Post*, (September 29, 1979), p. C1.

255. "The Significance of the SAT Decline Has Been Considerably Exaggerated," *NAESP Communicator*, 1 (November 21, 1977), p. 1.

256. Simmons, Althea T. L. "Statement of Althea T. L. Simmons, Director, Washington Bureau of the National Association for the Advancement of Colored People before the Subcommittee on Elementary, Secondary and Vocational Education of the House Committee on Education and Labor on H.R. 3564 and H.R. 4949." Paper presented to the House of Representatives Subcommittee on Elementary, Secondary, and Vocational Education, n.d. 13 pp.

257. "Six States Look at Truth-in-Testing Bills," *ECS Legislative Review*, 10 (February 11, 1980), pp. 1-2.

258.0 Slack, Warner V. and Douglas Porter. "The Scholastic Aptitude Test: A Critical Appraisal," *Harvard Educational Review*, 50 (May 1980), pp. 154-175.

258.1 ________________________________. "Training, Validity, and the Issue of Aptitude: A Reply to Jackson," *Harvard Educational Review*, 50 (August 1980), pp. 392-401.

259. Solomon, Steve. "The Nader Report on ETS," *Testing Digest*, (Spring 1980), p. 9.

260. ____________. "Test Coaching: The Official and Unofficial FTC Reports," *Testing Digest*, (Spring 1980), pp. 3-4.

260.1 "Special Test Dates Increased for New York Students," *Education Daily*, 13 (July 22, 1980), p. 1.

261. "Statement of Principles in Admissions and Recruitment," *Chronicle of Higher Education*, 18 (July 16, 1979), p. 5.

262. State of New York, Senate and Assembly. *An Act to Amend the Education Law, in Relation to Standardized Testing*. S. 5200-A/A. 7668-A, Cal. No. 1215, April 26, 1979. 3 pp.

263. State of New York, Senate. *An Act to Amend the Education Law, in Relation to Special Administrations of Standardized Tests*. 7537-A, February 6, 1980. 1 p.

264. State of New York, Senate. *An Act to Amend the Education Law, in Relation to Standardized Testing, and to Repeal Sections Three Hundred Forty-Two and Three Hundred Forty-Three of Such Law Relating Thereto*. 9142-A, Cal. No. 751, April 21, 1980. 4 pp.

265. Sternglass, Ernest. "The Nuclear Radiation/SAT Decline Connection," *Phi Delta Kappan*, 61 (November 1979), pp. 184-187.

266. Strenio, Andrew. "Excerpts from the Debate Over Open Versus Secure Testing: A Critical Review," *School Administrator*, 37 (January 1980), p. 13.

267. "Strong Academic Program Main Factor in Stable SAT Scores, Principals Say," *Education Daily*, 11 (April 5, 1978), p. 4.

268. "Students Get Way to Check Their Scores on SATs," *Washington Post*, (April 8, 1980), p. A5.

269. "Students and Learning," *Education Summary*, (December 1, 1979), pp. 6-7.

270. "Students May Check Scores on SATs; NEA Calls for Truth-in-Testing," *NEA Now*, (April 14, 1980), p. 2.

271. "Students More Materialistic, National Survey Shows." Los Angeles, CA: Office of Public Information, University of California, Los Angeles (405 Hilgard Ave., 90024), January 20, 1980. 4 pp.

272. "Study Links Score Drop to Baby Boom," *NASSP Newsletter*, 26 (December 1978), p. 2.

273. "Test Criticisms Are 'Open Assault' on College Admissions, Hanford Says," *Education Daily*, (February 8, 1980), p. 4.

274. "Test Scores Not Tops in Admissions Decisions, Study Finds," *Education Daily*, 12 (November 1, 1979), p. 5.

275. "Testing Agencies Urge Caution on FED Testing Bills," *Education Daily*, 12 (August 2, 1979), p. 2.

276. "Testing Service Wins $15,000 in Copyright Suit," *Chronicle of Higher Education*, 18 (March 26, 1979), p. 2.

277. Thomson, Scott D. "When a Criticism Is Not a Criticism," *Phi Delta Kappan*, 60 (March 1979), pp. 490-491.

278. _______________ and Nancy DeLeonibus. *Guidelines for Improving SAT Scores.* Reston, VA: National Association of Secondary School Principals (1904 Association Dr., 22091), 1978. 48 pp. ($3.00)

279. "Too Much Truth in Testing," *Washington Post*, (October 27, 1979), p. A12.

280. "Truth-in-Testing Bill Debate Begins Anew," *Education Daily*, 13 (June 6, 1980), p. 2.

281. "Truth-in-Testing Bills Introduced in Six States," *Education U.S.A.*, 22 (February 25, 1980), p. 194.

281.1 "Turnbull Resigns Presidency of ETS," *Education Daily*, 13 (October 27, 1980), p. 3.

282. Turnbull, William W. Letter to the editors. *New York Times*, (January 4, 1980.)

283. _________________. "A Partial View of ETS," *Phi Delta Kappan*, 59 (October 1977), pp. 103-106.

284. _________________. "Response to Remarks of Ralph Nader About Educational Testing Service at the Meeting of the American Psychological Association September 3, 1976." Princeton, NJ: Educational Testing Service, September 4, 1976. 6 pp.

285. _________________. "Statement by William Turnbull, President, Educational Testing Service," January 14, 1980. 3 pp.

286. "Two Researchers Debunk SATs as Aptitude Tests," *Washington Post*, (May 17, 1980), p. A2.

287. United States Congress. *A Bill to Require Certain Information Be provided to Individuals Who Take Standardized Educational Admissions Tests, and for Other Purposes.* H.R. 4949, 96th Congress, July 24, 1979. 15 pp.

288. "Updating Test Legislation," *ETS Developments*, 26 (1980), pp. 1, 10.

289.0 Watkins, Beverly T. "Truth-in-Testing Bills Dead till Spring," *Chronicle of Higher Education*, 19 (October 29, 1979), p. 13.

289.1 _________________. "Verbal Average Falls to 424, Math to 466 as Decline in SAT Scores Persists for 13th Year," *Chronicle of Higher Education*, 21 (October 6, 1980), p. 4.

290. Weaver, W. Timothy. "Educators in Supply and Demand: Effects on Quality," *School Review*, (August 1978), pp. 552-593.

291. _________________. "In Search of Quality: The Need for Talent in Teaching," *Phi Delta Kappan*, (September 1979), pp. 29-32, 46.

292. Weiss, Ted. "Statement by Representative Ted Weiss on the Educational Testing Act." Paper presented to the U.S. House of Representatives, July 24, 1979. 4 pp.

293. "When Your Job Is on the Line, Boost S.A.T. Scores with Coaching and Error Checks," *Executive Educator*, 1 (May 1979), pp. 20-22.

294. "Will Truth-in-Testing Break Up Testing Industry?" *Education U.S.A.*, 22 (January 7, 1980), pp. 137, 144.

295. "The Wirtz Panel Report: Good Writing," *Council for Basic Education Bulletin*, 22 (October 1977), p. 5.

296. Wirtz, Willard. *What Shall We Do About Declining Test Scores?* Washington, DC: American Association for Higher Education (One Dupont Circle, Suite 780, 20036), 1978. 8 pp. ($1.00)

297. Wirtz, Willard and Others. *On Further Examination.* Princeton, NJ: College Entrance Examination Board (Box 2815, 08540), 1977. 75 pp. ($4.00)

298. "Women and Math: Anxiety Discounted," *Guidepost*, 21 (October 5, 1978), pp. 7, 11. (A publication of the American Personnel and Guidance Association, Washington, DC)

299. Zais, Robert S. "The Decline of Academic Performance in the Classroom and the Reading Scores of Prospective Teachers: Some Observations," *High School Journal*, 62 (November 1978), pp. 52-57.

300. Zonana, Victor F. "Who Gets Ahead?" *Wall Street Journal*, (February 28, 1978), pp. 1, 18.